SÉRIE A, N° 3677
N° D'ORDRE : 4529

THÈSES

PRÉSENTÉES

A LA FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE PARIS

POUR OBTENIR

LE GRADE DE DOCTEUR ÈS SCIENCES MATHÉMATIQUES

PAR

JACQUES AGARD

Première thèse

ANALYSE STATISTIQUE ET PROBABILISTE DES SISMOGRAMMES

Deuxième thèse

PROPOSITIONS DONNÉES PAR LA FACULTÉ

soutenues le 17 avril 1961 devant la commission d'examen

MM. FORTET..... *Président.*
VILLE ..... } *Examinateurs.*
NEVEU ..... }

PUBLIÉ PAR

LA REVUE DE L'INSTITUT FRANÇAIS DU PÉTROLE
ET ANNALES DES COMBUSTIBLES LIQUIDES
Vol. XVI, 1961, n° 3, p. 1 à 48 et n° 4, p. 49 à 85.

—

PARIS

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41538123_0025

SÉRIE A, N° 3677
N° D'ORDRE : 4529

THÈSES

PRÉSENTÉES

A LA FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE PARIS

POUR OBTENIR

LE GRADE DE DOCTEUR ÈS SCIENCES MATHÉMATIQUES

PAR

JACQUES AGARD

Première thèse

ANALYSE STATISTIQUE ET PROBABILISTE DES SISMOGRAMMES

Deuxième thèse

PROPOSITIONS DONNÉES PAR LA FACULTÉ

soutenues le 17 avril 1961 devant la commission d'examen

MM. FORTET ..... *Président.*
VILLE ..... } *Examinateurs.*
NEVEU ..... }

PUBLIÉ PAR

LA REVUE DE L'INSTITUT FRANÇAIS DU PÉTROLE
ET ANNALES DES COMBUSTIBLES LIQUIDES
Vol. XVI, 1961, n° 3, p. 1 à 48 et n° 4, p. 49 à 85.

—

PARIS

FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE PARIS

Doyen..... J. PÉRÈS.

PROFESSEURS

G. JULIA	T	Analyse supérieure et Algèbre supérieure.
DE BROGLIE	T	Théories physiques.
PRENANT	T	Anatomie et Histologie comparées.
PÉRÈS	T	Mécanique des fluides et applications.
J. BOURCART	T	Géographie physique et Géologie dynamique.
PLANTEFOL	T	Botanique.
GRASSÉ	T	Evolution des êtres organisés.
PRÉVOST	T	Chimie organique.
BOULIGAND	T	Analyse appliquée à la Géométrie.
CHAUDRON	T	Chimie appliquée.
WYART	T	Minéralogie et Cristallogr.
TEISSIER	T	Zoologie.
MANGENOT	T	Biologie végétale (Orsay).
P. AUGER	T	Physique quantique et relativité.
MONNIER	T	Physiologie des fonctions.
PIVETEAU	T	Paléontologie.
ROCARD	T	Physique (E. N. S.).
H. CARTAN	T	Mathématiques (E. N. S.).
LAFFITTE		Chimie générale.
FAVARD	T	Géométrie supérieure.
COULOMB	T	Physique du globe.
Mlle COUSIN	T	Biologie.
CHRÉTIEN	T	Chimie minérale.
P. DRACH	T	Zoologie.
KASTLER	T	Physique (E. N. S.).
EPHRUSSI	T	Génétique.
RIVIÈRE	T	Géologie (Orsay).
GAUTHERET	T	Biologie (P. C. B.).
R. LUCAS	T	Recherches physiques.
A. THOMAS	T	Biologie cellulaire.
ARNULF	T	Optique appliquée.
MAX MORAND	T	Physique Enseignement.
SOLEILLET	T	Physique (P.C.B.).
FORTIER	T	Mécanique expérimentale des fluides.
DANJON	T	Astronomie.
PETIT	T	Biologie maritime.
QUENEY	T	Météorologie et dynam. atmosphérique.
GALLIEN	T	Embryologie.
EICHHORN	T	Biologie végét. (P. C. B.).
DE CUGNAC	T	Biologie végét. (S. P. C. N.).
Mlle CAUCHOIS	T	Chimie physique.
THELLIER	T	Physique du globe.
L'HÉRITIER	T	Biologie générale (Orsay).
GUIVET	T	Radio-électricité (Orsay).
PONCIN	T	Mécanique générale.
DUBREIL	T	Arithmétique et théorie des nombres.
QUELET	T	Chimie organique.
CAGNIARD	T	Géophysique appliquée.
CHAMPETIER	T	Chimie macromoléculaire.
CUVILLIER	T	Micropaléontologie.
JUNG	T	Pétrographie.
TRILLAT	T	Microscopie et diffraction électronique.
WIEMANN	T	Chimie organ. et structur.
JACQUINOT	T	Spectroscopie (Orsay).
VASSY	T	Physique de l'atmosphère.
DESTOUCHES	T	Théories physiques.
P. PRUVOST	T	Géologie.
AMIEL	T	Chimie des solides.
HOCART	T	Minéralogie et Cristallogr.
J.-P. MATHIEU	T	Recherches physiques.
COUTEAUX	T	Biologie animale (P. C. B.).
MAY	T	Zoologie (Orsay).
CHOQUET	T	Théorie des fonct. et Topologie.
FELDMANN	T	Biologie végétale (P. C. B.).
GUINIER	T	Physique des solides (Orsay).
JOST	T	Phy-iologie comparée
FORTET	T	Calcul des probabilités.
SCHWARTZ	T	Calcul différent. et intégral.
CHOUARD	T	Physiologie végétale.
MALAVARD	T	Aviation (Techn. aéronaut.).
BRELOT	T	Calcul différentiel
NORMANT	T	Synthèse organ. générale.
BÉNARD	T	Chimie (E.N.S.C.P.).
BUVAT	T	Botanique (E. N. S.).
DUGUÉ	T	Statistiques mathématiques.
SOULAIRAC	T	Psychophysiologie.
ULRICH	T	Physiologie vég. (E.N.S.).
MARÉCHAL	T	Optique théorique et appliquée.
KIRRMANN	T	Theories chimiques.
CHADEFAUD	T	Botanique.
Mlle LE BRETON	T	Physiologie.
SALEM	T	Mathématiques générales.
LELONG	T	Mathématiques générales.
DEVILLERS	T	Anatomie et Histologie comparées.
EHRESMANN	T	Topologie algébrique.
FRANÇON	T	Physique (S. P. C. N.).
GLANGEAUD	T	Géogr. physique et Géologie dynamique.
GODEMENT	T	Mathématiques (M. P. C.).
PISOT	T	Calcul des probabilités.
ROCH	T	Geologie
SCHATZMANN	T	Astrophysique.
TERMIER	T	Paléontol. stratigraphique.
ZAMANSKY	T	Mathématiques générales.
LENNUIER	T	Physique (P. C. B.).
RIZET	T	Génétique (Orsay).
ROUTHIER	T	Géologie appliquée.
Mme TONNELAT	T	Relativité.
DIXMIER	T	Mathémat ques (M. P. C.).
SOUCHAY	T	Chimie (P.C.B.).
AIGRAIN	T	Electrotechnique générale.
BRUSSET		Chimie minérale.
M. LÉVY	T	Physique théorique des hautes énergies (Orsay).
Mme CHAIX	T	Chimie biologique.
Mme HUREL-PY	T	Biologie végétale (P. C. B.).
PIAUX	T	Chimie (S. P. C.N.).
BRUN	T	Mécan. expér. des fluides.
LEDERER	T	Chimie biologique.
Mme DUBREIL	T	Mathématiques Agreg.).
Mme LELONG	T	Mathématiques (E. N. S.).
BELLAIR	T	Géologie (S. P. C. N.).
COTTE	T	Physique (M. P.C.).
J.-E. DUBOIS	T	Chimie (M.P.C.).
LAMOTTE	T	Zoologie (E. N. S.).
A. MICHEL	T	Chimie minérale (Orsay).
OLMER	T	Energétique générale.
ROUAULT	T	Physique.
GAUTHIER	T	Mécanique appliquée.
BARCHEWITZ	T	Chimie physique.
BROSSEL	T	Physique atomique.
BUSER		Biologie animale (P.C.B.).
CAMUS		Biologie végétale (P.C.B.).
CASTAING	T	Physique (Orsay).
CURIEN	T	Minéralogie et Cristallogr.
MOYSE	T	Phy-iologie vég. (Orsay)
PANNETIER	T	Chimie (Orsay).
POSSOMPES	T	Biologie animale (P.C.B.).
PULLMANN		Chimie théorique.
TEILLAC	T	Physique nucléaire et Radioactivité. (Orsay).
TONNELAT		Biologie physico-chimique.
VILLE	T	Econométrie
VILLEMART		Chimie (P. C. B.).
DODÉ	T	Chimie (Orsay).
FREYMANN	T	Physique (S.P.C.N.).
GUINOCHET	T	Biologie végétale (Orsay).
ROLLET	T	Chimie (P. C. B.).
Mlle JOSIEN	T	Chimie (P. C. B.).
CHEVALLEY	T	Géométrie algébrique et théorie des groupes.
ARNOULT	T	Électronique (Orsay).
CHAPELLE	T	Physique (Orsay).
DELANGE	T	Mathematiques (Orsay).
DENY	T	Mathém. gener. (Orsay).
P. GERMAIN	T	Mec. théorique des fluides.
G. LUCAS	T	Géologie.
ALLARD	T	Chimie physique.
BERTHELOT	T	Phys. fondam. des hautes énergies (Orsay).
BRICARD	T	Météorologie.
Mme ALBE-FESSARD		Psychophysiologie.
FRIEDEL		Phys. des solides (Orsay).
M. JULIA		Chimie organique structur.
LENDER		Biologie animale (S.P.C.N.) (Orsay).
MAGAT		Physico-chimie des Rayonn.
Mlle QUINTIN	T	Electrochimie
MONOT	T	Biochimie cellulaire.
BENOIT	T	Physiologie anim. (Orsay).
DE POSSEL	T	Analyse numérique.
CHARLOT	T	Chimie analytique.
HAISSINSKY	T	Radiochimie.
LEMÉE	T	Physiol. (B.M.P.V.). (Orsay)
GUÉRIN	T	Chimie (Orsay).
M. CHATELET		Chimie (P. C. B.).
JEAN		Physique (Orsay).
MATTLER	T	Physique (Orsay).
L. MICHEL		Physique théor. (Orsay).
RUMPF		Chimie (P. C. B.).
BERGERARD		Biologie anim. (S. P. C. N.) (Orsay).
BERTEIN		Électronique (Orsay).
BOUISSIÈRES		Chimie des Radioéléments (Orsay).
Mme COUTURE		Thermodynamique et Mécanique physique.
ELLENBERGER		Géologie appliquée.
FRANC		Biologie animale (P. C. B).
SCHNELL		Botanique tropicale (Orsay).
STOLKOWSKI		Biologie anim (P.C.B.).
ACHER		Chimie biol. (Orsay).
BAUDOIN		Biologie.
BEAUMONT		Biologie animale.
BLAMONT		Géophysique Ionosphère.
BOURREAU		Biologie végét. (S. P. C. N.) (Orsay).
CURIE D		Physique (M.G.P.).
DURAND-DELGA		Géologie.
FRÉON		Chimie organique (Orsay).
HELLER		Physiologie végétale
TORTRAT		Calcul des Probabilités.
CAILLEUX		Géologie (S. P.C.N.).
MAGNAN	T	Physique (P.C.B.).
DAUDEL		Mécanique ondulatoire appliquée à la chimie théor. et à la Physique nucléaire.
Mlle FOURCROY		Biologie végét (S.P.C.N.).
BARRAUD		Physiologie cellulaire.
NOIROT		Evolut. des êtres organisés.
Mme FOURÈS-BRUHAT	T	Mécanique analytique et Mécanique céleste.
Mlle VEIL		Physiologie générale.
LWOFF	T	Microbiologie.
SIESTRUNCK	T	Mécanique physique et expérimentale.
LACOMBE	T	Métallurgie (Orsay).
BROCHARD	T	Physique (P.C.B.) (Orsay).
CABANNES	T	Mécanique générale.
LEFEBVRE	T	Mathém. (M P.C.) (Orsay).
MAZET	T	Mécanique générale (Orsay).

Secrétaire général..... R. POULAIN

R. UNIV BIBLIOTHEEK LEIDEN

Nous exprimons toute notre reconnaissance à M. le professeur FORTET qui a guidé et suivi de près l'avancement progressif de nos travaux.

Nous remercions également l'*Institut Français du Pétrole* qui fut le promoteur de cette étude, et tout particulièrement MM. CHOLET et GRAU qui nous ont apporté leurs conseils techniques, leurs idées et une aide matérielle considérable.

Enfin, nous remercions les ingénieurs et les assistants de l'*Institut Français du Pétrole* qui ont participé à l'élaboration des calculs contenus dans cette thèse, et parmi eux MM. BOIS, FAIL, HEMON et SCHWING.

SOMMAIRE

INTRODUCTION

I. DESCRIPTION DE LA PROSPECTION SISMIQUE

1. Enregistrement d'un sismogramme 6
2. Aspect aléatoire des sismogrammes 7
3. Problèmes posés par les géologues 7
4. Formulation mathématique du problème 7

II. ANALYSE DES COMPOSANTES DU SIGNAL $y(t)$

1. Étude des vitesses de propagation 8
2. Étude des coefficients de réflexion 12
3. Étude de l'impulsion $s(t)$ 26
4. Étude des bruits 26
5. Étude des amplitudes des sismogrammes 26

III. ÉTUDE STATISTIQUE DES AMPLITUDES

1. Étude d'un cas simple $y(t) = s(t - t_0) + \varepsilon(t)$ 30
2. Étude des procédés de mixage 32

IV. NOMBRE DE RÉFLEXIONS, FILES D'ATTENTE, MULTIPLES

1. Nombre de réflexions en t 33
2. Réflexions multiples 35

V. ANALYSE HARMONIQUE, AUTOCORRÉLATION ET INTERCORRÉLATION

1. Intérêt de ces calculs 39
2. Analyse harmonique 39
3. Autocorrélation 41
4. Variance de l'autocorrélation 44
5. Intercorrélation 45
6. Analyse d'une composante périodique 46

VI. CORRÉLATION MOBILE ET FILTRAGE, CONTROLE DE GAIN

1. Corrélation mobile 49
2. Contrôle de gain 52

VII. DÉTERMINATION DE S (T) PAR L'AUTOCORRÉLATION

1. Méthode de Robinson 55
2. Méthode de Horton 55
3. Assimilation de $s(t)$ à une période de sinusoïde 57
4. Essai d'une sinusoïde amortie $y(x) = e^{-\alpha x} \sin \beta x$ 58
5. Détermination d'un signal de Ricker 59
6. Méthode de Lindsey : élimination des réflexions fantômes 59

VIII. MÉTHODE DES SISMOGRAMMES SYNTHÉTIQUES

1. Calcul du sismogramme 62
2. Autocorrélation et corrélation mobile 65
3. Élimination du bruit 66
4. Fabrication de coefficients de réflexion synthétiques 67

IX. ÉLIMINATION DES BRUITS A L'ENREGISTREMENT

1. Étude des traces de référence 68
2. Résultats de l'expérimentation 73
3. Analyse des amplitudes 79

X. DÉTECTION AUTOMATIQUE DES RÉFLEXIONS

1. Caractéristiques d'une réflexion 82
2. Probabilité d'apparition d'une réflexion 82
3. Probabilité d'une fausse réflexion 83
4. Distribution de probabilité des réflexions 83
5. Détection automatique des réflexions 84

BIBLIOGRAPHIE

INTRODUCTION

Pour connaître la constitution géologique d'une région, les géologues procèdent à diverses mesures de surface. Pour préciser en particulier les courbes de niveau des couches géologiques, ils emploient les méthodes de la prospection sismique : une explosion est provoquée au voisinage des couches superficielles du terrain étudié. L'onde de choc de l'explosion se propage dans le sous-sol. Les couches géologiques étant de densité et de nature différentes, lorsque le front de l'onde franchit la frontière séparant deux couches, une partie de l'énergie transportée est réfléchie vers la surface du fait des discontinuités des constantes élastiques des couches. En surface, des sismographes enregistrent les arrivées de l'énergie libérée par l'explosion durant les deux ou trois secondes qui la suivent. En pointant ces arrivées sur les sismogrammes et en recoupant leurs indications avec celles procurées par d'autres mesures, le géologue précise sa connaissance du sous-sol : position des discontinuités, vitesses de propagation, coefficients d'absorption... Il peut ainsi décider de l'intérêt de sondages ultérieurs.

L'importance de l'enjeu des recherches géologiques et le coût très élevé de la prospection sismique effectuée en dernier lieu avant les sondages nécessitent une exploitation complète des informations contenues dans les sismogrammes. Malheureusement, le pointé des couches réfléchissantes est souvent rendu impossible par l'existence de bruits de natures diverses, ou par la superposition d'ondes de surface. De plus, le contenu des sismogrammes est assimilable à un signal aléatoire puisque la forme et l'emplacement des signaux réfléchis sont a priori inconnus et que des bruits divers s'y ajoutent. C'est pourquoi l'analyse statistique et probabiliste a été appliquée de longue date en sismique, tout spécialement lorsque le brouillage rend les sismogrammes inexploitables par une lecture directe.

Le but de cette thèse est l'application, à l'analyse des sismogrammes, des méthodes statistiques et probabilistes. Ces méthodes facilitent inégalement la détection des réflexions. Des idées théoriquement séduisantes se révèlent inapplicables en pratique. Dans bien des cas, le rapport signal/bruit est trop faible et les spectres du signal et du bruit s'interpénètrent trop pour qu'il reste un espoir d'extraire le signal du bruit. Les hypothèses simples nécessaires à l'élaboration de la théorie ne sont pas toujours vérifiées.

Notre contribution a consisté à essayer les méthodes déjà proposées dans le domaine de la sismique et à compléter l'analyse physique des phénomènes : lois de répartition des vitesses de propagation du son dans le sous-sol, lois de répartition des coefficients de réflexion, vérification de leur répartition à la Poisson, étude du bruit et de son élimination. Les procédés d'analyse harmonique, d'autocorrélation, d'intercorrélation, de corrélation mobile et de filtrage sont discutés avec des exemples à l'appui. Certaines méthodes de détermination du signal engendré par l'explosion sont analysées.

Les procédés statistiques étudiés ici peuvent probablement être transposés à d'autres séries temporelles que les sismogrammes. Nous faisons même à propos du comptage des réflexions une digression sur le problème des attentes où aux réflexions correspondent des arrivées de clients et où les impulsions $s(t)$ sont remplacées par une fonction qui reste égale à 1 durant la durée aléatoire du temps de service du client.

Enfin, nous avons utilisé des « sismogrammes synthétiques ». Il s'agit de sismogrammes construits, comme les sismogrammes réels, à partir de coefficients de réflexion connus par des sondages et filtrés par diverses impulsions $s(t)$. Ils peuvent être mélangés à des bruits de natures diverses.

L'emploi de ces sismogrammes synthétiques permet de tester les méthodes statistiques proposées sur des exemples communs offrant une gamme importante de combinaisons. Connaissant les constituants de ces sismogrammes, on peut analyser la valeur d'une méthode pour extraire ces constituants.

Cette thèse contient peu de résultats mathématiques nouveaux. Nous sommes très peu sorti du cadre réel du problème posé pour résoudre des problèmes connexes moyennant des hypothèses simplificatrices mais irréelles. Les conditions dans lesquelles se pose le problème de la détection sont rebelles au traitement mathématique. Certaines hypothèses de stationnarité sont déjà mal satisfaites. Il n'était pas question de simplifier davantage sans s'écarter du réel.

Il faut enfin noter que les variances des « meilleures » estimations sont fondamentales. Le calcul de ces variances est malheureusement très délicat. C'est pourquoi nous avons souvent utilisé des méthodes simples ou des procédés de simulation, en laissant posée la résolution mathématique pure.

I. DESCRIPTION DE LA PROSPECTION SISMIQUE

Sans entrer dans le détail des méthodes sismiques, nous en donnons ici une brève description : on observe les ondes élastiques qui reviennent à la surface du sol après s'être propagées dans le sous-sol à la suite d'un ébranlement. Dans le sous-sol, les discontinuités des constantes électriques ou des densités entraînent des réflexions, réfractions ou diffractions de l'onde provenant de l'ébranlement. Leur étude doit nous révéler la position des discontinuités, les vitesses de propagation et les coefficients d'absorption.

1. Enregistrement d'un sismogramme (fig. 1).

Généralement, on procède à une explosion de dynamite dans les couches superficielles du terrain. Des sismographes disposés en surface sont reliés par des câbles électriques à des appareils amplificateurs et enregistreurs à plusieurs canaux. Dans les deux ou trois secondes qui suivent un tir,

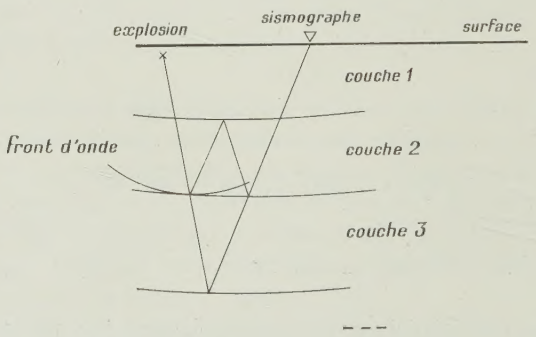


FIG. 1.

on enregistre sur papier photographique ou sur bande magnétique les retours d'énergie de l'explosion. Les groupes de sismographes disposés en ligne par rapport au lieu de l'explosion donnent naissance à des traces (de 12 à 24 par sismogramme). On applique ensuite à ces traces des corrections de temps correspondant aux différences des temps de propagation du point de tir aux sismographes,

ainsi qu'aux différences éventuelles d'altitude des sismographes (exemple de sismogramme sur la figure 21, page 61).

La détection des réflexions est ensuite faite par des spécialistes qui identifient à vue les réflexions sur l'enregistrement photographique du sismogramme par des analogies d'amplitude, de forme et de mise en phase constatées sur toutes les traces et permettant même de suivre le profil d'une couche sur les sismogrammes des tirs voisins.

Nous n'insisterons pas davantage sur l'aspect pratique des méthodes sismiques. On trouvera dans la bibliographie une liste d'ouvrages traitant la question (1) (2) (11).

2. Aspect aléatoire des sismogrammes.

Théoriquement, un sismogramme a un contenu certain : l'ensemble des réflexions consécutives à l'explosion, provenant de couches géologiques d'emplacement déterminé. Mais, a priori, ces emplacements sont inconnus. De plus, si l'on effectue deux tirs au même endroit, les deux sismogrammes diffèrent notablement, du fait de nombreux facteurs aléatoires. Les impulsions $s\,(t)$ engendrées par les explosions ne sont pas identiques. Elles subissent une déformation et un allongement avec la profondeur et l'absorption. Enfin, on peut discerner trois sortes de bruits :

Le bruit naturel dont la fréquence est un peu supérieure à 50 Hz est dû à l'agitation existant à la surface du sol. Il étouffe surtout les signaux faibles de la fin d'un enregistrement.

Le bruit provoqué par le tir contient des réflexions multiples, des diffractions et des ondes de surface (*groundroll*).

Des oscillations parasites dues à un mauvais couplage du sismographe au sol, ou des réglages mal faits, ainsi que le bruit de l'appareillage brouillent enfin le signal.

3. Problèmes posés par les géologues.

Les sismiciens désirent éliminer les bruits naturels et les oscillations parasites, ils veulent identifier les bruits provoqués par le tir pour en tirer une information et en tenir compte dans l'identification des réflexions qui devrait être faite par des procédés automatiques.

4. Formulation mathématique du problème.

Si nous supposons connue la courbe $k\left(\frac{t}{2}\right)$ des coefficients de réflexion qui permettent d'enregistrer en t une impulsion émise à l'instant 0 et si $s\,(t)$ est l'impulsion engendrée par l'explosion, on peut dire que l'amplitude du signal résultant des réflexions simples revenant en surface en t est :

$$y\,(t) = \int_{t-\mathfrak{T}}^{t} k\left(\frac{\tau}{2}\right) s\,(t-\tau)\,d\tau$$

où $\mathfrak{T}$ est la longueur de l'impulsion $s\,(t)$.

Il convient d'ajouter les bruits divers $\varepsilon\,(t)$ et les réflexions multiples ainsi que les bruits périodiques $g\,(t)$.

En réalité, comme l'énergie de l'explosion est absorbée lors de sa diffusion, l'énergie reçue au cours de l'enregistrement décroît presque exponentiellement et on procède à un contrôle de gain, automatique ou programmé, qui a pour but de maintenir l'énergie enregistrée presque constante. Comme l'impulsion $s\,(t)$ est courte (50 à 100 ms devant les deux à trois secondes d'enregistrement),

on supposera que, du fait du contrôle de gain, la fonction $k\,(t)$ est stationnaire et que $s\,(t)$ n'est pas modifié. On aura :

$$y\,(t) = \int_{t-\mathfrak{T}}^{t} k\left(\frac{\tau}{2}\right) s\,(t-\tau)\,d\tau + \varepsilon\,(t) + g\,(t)$$

Si nous désirons écrire l'équation exacte du signal enregistré, il convient en fait de tenir compte de l'influence des réflexions multiples revenant en surface à l'instant t. Pour cela, il faut remplacer dans l'équation précédente $k\left(\frac{t}{2}\right)$ par la somme de $k\left(\frac{t}{2}\right)$ plus la contribution de tous les multiples. GRAU a mis au point une méthode permettant le calcul effectif de tous les multiples à partir de la donnée de $k\left(\frac{t}{2}\right)$. Voir en (2) et (8).

II. ANALYSE DES COMPOSANTES DU SIGNAL $y\,(t)$

Avant d'examiner les méthodes statistiques permettant de retrouver la courbe $k\,(t)$ à partir de $y\,(t)$ nous allons étudier la nature statistique des composantes de $y\,(t)$.

1. Étude des vitesses de propagation.

Lorsqu'on procède à un sondage par forage, on fait parfois des mesures des vitesses de propagation du son aux diverses profondeurs. On obtient ainsi des « carottages continus » de vitesse. Ils montrent que les vitesses varient très vite autour de leurs valeurs moyennes. Ces variations peuvent être dues à des hétérogénéités locales. Seules les grandes variations de vitesse correspondant à un changement de couche géologique se suivent sur toute une région.

En calculant le temps mis par une impulsion pour venir se réfléchir à une profondeur h et revenir en surface, on peut faire correspondre aux temps t de retour des réflexions les vitesses $V\left(\frac{t}{2}\right)$ des couches réfléchissantes correspondantes. On obtient ainsi la courbe de vitesse $V\left(\frac{t}{2}\right)$ indiquée dans la figure 2.

a. Distribution des vitesses

Il semble que les vitesses évoluent avec des fluctuations aléatoires sans corrélation autour de certains paliers. A certains instants, on enregistre des sauts brutaux des vitesses qui évoluent ensuite autour d'un nouveau palier. On trouvera sur la figure 3 des ajustements à la loi normale de quatre distributions expérimentales de vitesses. On peut admettre que ces ajustements ne s'écartent pas trop de la droite de Henri. Certains comme PUISELET et GRÉOUX sont peu satisfaisants parce qu'ils sont composés de plusieurs paliers de vitesses et que les vitesses ne sont pas stationnaires le long de ces paliers. Si l'on excepte le cas de GRÉOUX, les vitesses évoluent autour de 3 000 m/s et l'écart type est voisin de 800 m/s. L'ajustement normal est d'autant plus approximatif que les vitesses sont positives, ce qui nécessiterait de tronquer les lois normales théoriques pour lesquelles des vitesses négatives ont des probabilités non nulles.

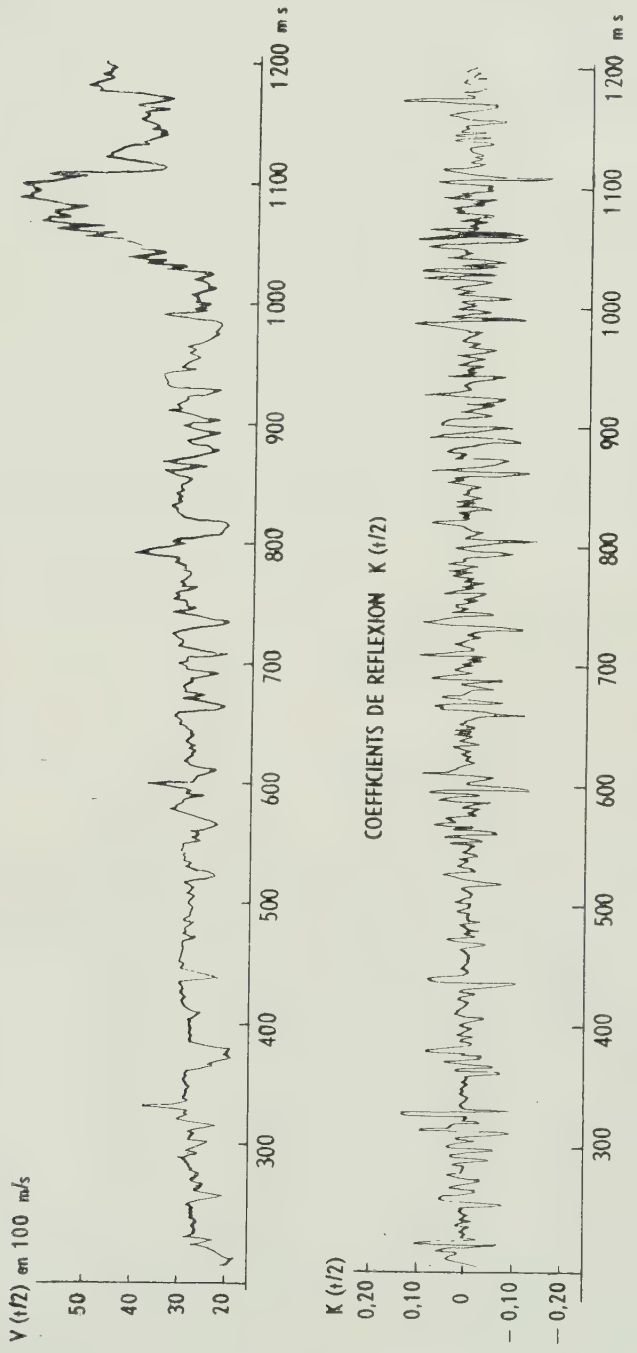

VITESSE DE PROPAGATION DU SON
V (t/2) en 100 m/s
50
40
30
20
300
400
500
600
700
800
900
1000
1100
1200 ms
COEFFICIENTS DE REFLEXION K (t/2)
K (t/2)
0,20
0,10
0
− 0,10
− 0,20
300
400
500
600
700
800
900
1000
1100
1200 ms

FIG. 2.

b. Espacements entre $|\Delta V| >$ seuil.

Comme a priori il n'y a pas de raison de penser que les frontières des couches géologiques obéissent à une répartition particulière, on peut penser qu'elles sont réparties à la Poisson. C'est ce que nous avons pu vérifier sur plusieurs exemples (2). La figure 4 donnée à titre d'exemple montre que la loi régissant la distance t entre deux instants consécutifs t_i et t_{i+1} où les variations de vitesse $|V(t_i) - V(t_{i+dt})|$ ont dépassé un seuil $|\Delta V|$ fixé est la loi exponentielle :

$$\text{probabilité d'une distance} > t = e^{-\alpha t}$$

Cette figure a été tracée sur un papier semi-logarithmique. Sur la figure 4 figurent trois ajustements correspondant à des méthodes de lecture différentes. Le relevé des vitesses pose en effet un problème

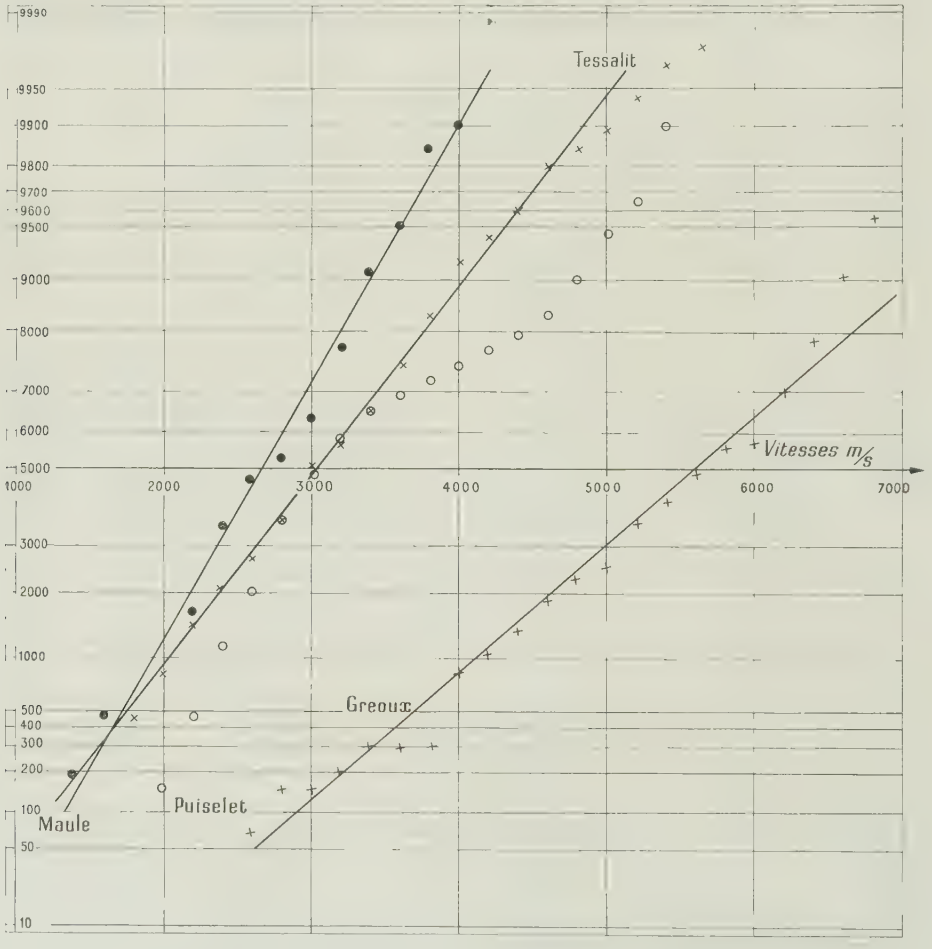


FIG. 3. — Ajustement à la loi normale des vitesses.

MAULE { $m = 2\,660$ m/s. $\sigma = 580$ » $n = 221$ observations.

TESSALIT { $m = 3\,040$ m/s. $\sigma = 800$ » $n = 959$ observations.

PUISELET { $m = 3\,050$ m/s. $\sigma = 800$ » $n = 533$ observations.

GREOUX { $m = 5\,600$ m/s. $\sigma = 1\,140$ » $n = 132$ observations.

d'échantillonnage. On peut relever les vitesses moyennes de l'instant t à $t + dt$ en estimant que les écarts à cette moyenne sur un intervalle dt sont négligeables. On peut également relever les vitesses exactes aux instants t, $t + dt$, $t + 2\,dt$... Les deux méthodes donnent des résultats quelque peu différents. Prendre la valeur moyenne de V entre t et $t + dt$ revient à imposer un filtrage de plus au filtrage résultant de l'échantillonnage de période dt. La deuxième méthode paraît donc préférable et d'un relevé plus simple. On voit sur la figure 4 que les ajustements graphiques sont assez médiocres

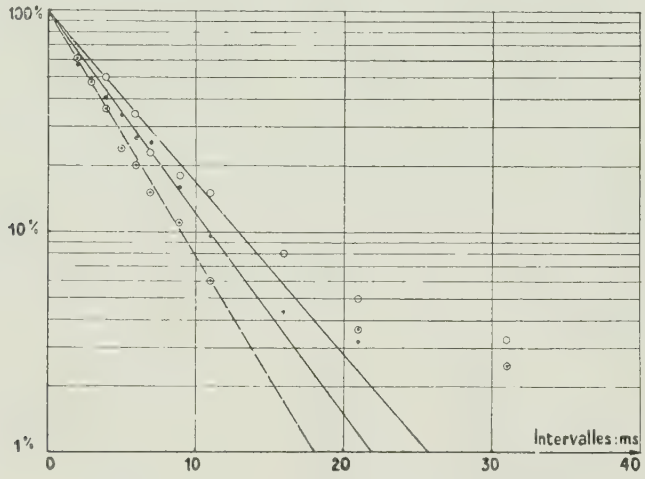


FIG. 4. — Ajustement exponentiel.

Intervalles entre vitesses.
$|\Delta V| > 300$ m/s
● 93 obs. $\bar{V}$, $dt = 1$ ms.
⊙ 80 » V, $dt =$ »
○ 61 » V, $dt = 2$ ms.

et qu'il y a trop de grands intervalles. Le fait de moyenner les vitesses accroît naturellement le nombre des grands intervalles, car, en moyennant, on élimine les grands écarts de vitesse.

Enfin, si l'on accroît l'intervalle d'échantillonnage en passant de $dt = 1$ ms à $dt = 2$ ms, la distance moyenne entre les instants t tels que $|v(t) - v(t + dt)| > 300$ m/s est accrue.

Ces observations s'expliquent bien par la théorie. La vitesse $v(t)$ est supposée stationnaire et soit $f(v)$ sa loi avec une moyenne $\bar{v}$ et une autocorrélation $\rho(h)$ entre deux instants t et $t + h$.

La probabilité que $|v(t) - v(t + h)| > \Delta V$ est :

$$P(\Delta V) = \int_0^\infty \left[\int_0^{v_1 - \Delta V} f(v_1, v_2)\, dv_2 + \int_{v_1 + \Delta V}^\infty f(v_1, v_2)\, dv_2 \right] dv_1$$

avec $v_1 = v(t)$ et $v_2 = v(t + h)$.

Si $v(t)$ suit une loi normale, on aura :

$$P_h(\Delta V) = 2 \int_0^\infty \int_0^{v_1 - \Delta V} \frac{1}{2\pi\sigma^2\sqrt{1 - \rho^2(h)}}\, e^{-\frac{1}{2\sigma^2}\left[\frac{(v_1 - \bar{V})^2 - 2\rho(h)(v_1 - \bar{V})(v_2 - \bar{V}) + (v_2 - \bar{V})^2}{1 - \rho^2(h)}\right]} dv_2 dv_1$$

Si $h = 0$, $\rho(h) = 1$ et $P(\Delta V) = 0$, résultat évident.

Lorsque h croît, l'autocorrélation $\rho(h)$ décroît. Dans le cas d'une loi normale, la différence $v_1 - v_2$ suit une loi normale de moyenne nulle et d'écart type $\sigma\sqrt{2[1-\rho(h)]}$. Lorsque l'intervalle d'échantillonnage dt croît, la probabilité d'un écart absolu de vitesse supérieur à ΔV — identique quel que soit t d'après la stationnarité de $v(t)$ — croît jusqu'à la valeur h_0 où $\rho(h_0)$ passe par un minimum. Pour les valeurs de h où $\rho(h)$ tend vers zéro en espérance mathématique, $P(\Delta V)$ est constant. Sur un secteur de longueur T et avec un échantillonnage h, le nombre moyen d'instants t tels que $|v(t) - v(t+h)| > \Delta V$ est $\frac{T}{h} P_h(\Delta V)$. En passant d'un échantillonnage de 1 à 2 ms, on ne dédouble pas tout à fait le nombre de variations de vitesse $> \Delta V$ parce que $\rho(h)$ décroît un peu, ainsi qu'on peut le constater sur la figure 4.

Lorsqu'on considère $\bar{v}(t) = \frac{1}{h}\int_t^{t+h} v(\tau)\,d\tau$, le nombre moyen des instants où $|\bar{v}(t) - \bar{v}(t+h)| > \Delta V$ est plus faible que pour $|v(t) - v(t+h)|$ car la moyenne $\bar{v}(t)$ a un écart type inférieur à celui de $v(t)$ et de ce fait les grands écarts de vitesse entre t et $t+h$ sont plus rares, et les intervalles entre des instants où $|\bar{v}(t) - \bar{v}(t+h)| > \Delta V$ sont plus longs.

2. Étude des coefficients de réflexion.

a. Étude pratique.

A chaque profondeur, on peut associer un coefficient de réflexion le plus souvent très voisin de zéro. Pratiquement, on peut construire la courbe des coefficients de réflexion $k\left(\frac{t}{2}\right)$ (à partir de la courbe des vitesses $V(t)$) correspondant aux instants t de retour des réflexions ayant eu lieu à l'instant $\frac{t}{2}$ après l'explosion. On peut avoir une bonne approximation de $k\left(\frac{t}{2}\right)$ en adoptant un échantillonnage de $V\left(\frac{t}{2}\right)$ de 1 ou 2 ms, et en calculant :

$$k\left(\frac{t}{2}\right) = \frac{V\left(\frac{t}{2}\right) - V\left(\frac{t}{2} + dt\right)}{V\left(\frac{t}{2}\right) + V\left(\frac{t}{2} + dt\right)} \text{ (fig. 2)}$$

On néglige ainsi l'effet relativement peu important des variations de densité du sol S, puisque théoriquement on a :

$$dk = \frac{1}{2}\, d\log(SV)$$

Les difficultés signalées au sujet du relevé des $V\left(\frac{t}{2}\right)$ se retrouvent sur les valeurs résultantes de $k\left(\frac{t}{2}\right)$. La connaissance de la loi des $k\left(\frac{t}{2}\right)$ étant importante, nous avons étudié de nombreuses répartitions empiriques.

Distribution des coefficients de réflexion.

L'ajustement des coefficients $k(t)$ à la loi normale a été tenté sur plusieurs exemples (2). La figure 5 fournit un exemple d'ajustement. Les intervalles entre les $k(t)$ sont en principe de 1 ms. On a constaté que les ajustements ne sont pas satisfaisants dans l'ensemble. On distingue par

exemple sur la figure 5 trois alignements de points sur le graphique de Henri. Cela peut signifier que la loi de Gauss ne s'ajuste pas. D'autres lois (log-normale appliquée aux $|k|$, ...) ne donnent pas de meilleurs résultats. L'existence des trois alignements peut également provenir du mélange

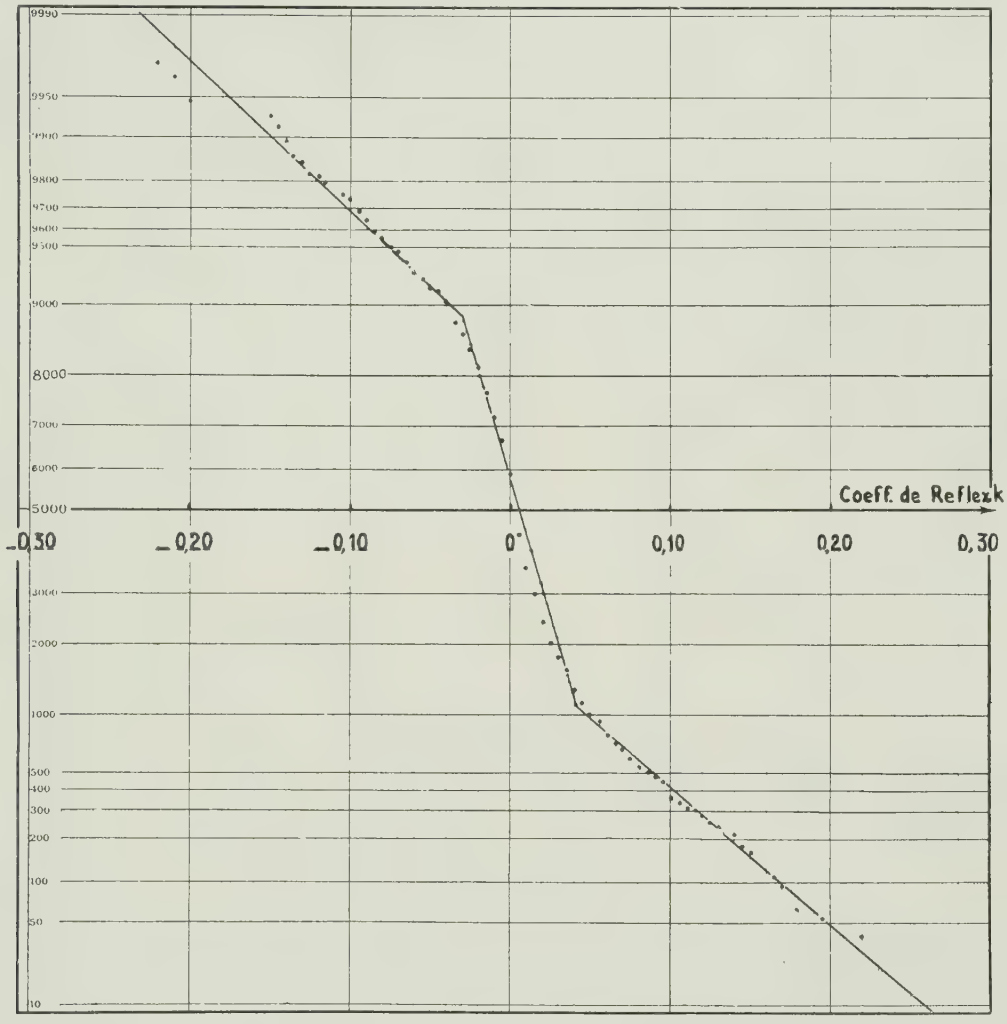


FIG. 5. — Ajustement à une loi normale des coefficients de réflexion.

Puits D.
● 734 observations.
k, dt = 1 ms.
moyenne = 0,005.
écart type = 0,11.

de deux populations. Ces deux populations pourraient provenir de deux couches lithologiques différentes pour lesquelles les écarts types des k (t) différeraient. L'une des populations peut également provenir d'erreurs de lecture. Le mode de calcul de k (t) revient à lire les valeurs de V (t) sur la courbe des vitesses, puis à calculer k (t) à partir de ces valeurs. Or, la lecture de V (t) introduit des erreurs approximativement proportionnelles à la vitesse. Lorsque les variations de vitesse entre t

et $t + h$ sont presque nulles, on peut lire des différences plus fortes par suite d'erreurs. On remplacerait ainsi de nombreux coefficients $k\,(t)$ nuls par des coefficients de moyenne nulle suivant une loi normale de faible écart type. Les coefficients de réflexion correspondant à des changements de couche géologique pourraient suivre à leur tour une loi normale de moyenne nulle mais dont la dispersion serait beaucoup plus forte. Plus l'échantillonnage est serré, plus la proportion des coefficients de réflexion correspondant à de vrais changements de nature de terrain décroît. La combinaison dans une proportion inconnue a priori de coefficients réels et de coefficients dus à des erreurs de lecture peut alors fournir trois alignements sur le graphique de Henri. Deux des alignements pratiquement parallèles correspondent à la population des coefficients réels à forte dispersion. L'alignement du centre correspond essentiellement à la population des coefficients des erreurs de lecture ou à la population de la couche lithologique à faible écart type.

La densité de la loi des $k\,(t)$ est $f\,(x)$, avec,
les populations 1 et 2 suivant des lois normales,
p_1 = proportion de la population 1,
$p_2 = 1 - p_1$ = proportion de la population 2,
σ_1 = dispersion de la population 1,
σ_2 = dispersion de la population 2,
les moyennes sont nulles ou égales [1] :

$$f(x) = \frac{1}{\sqrt{2\pi}}\left[\frac{p_1}{\sigma_1}\, e^{-\frac{x^2}{2\sigma_1^2}} + \frac{p_2}{\sigma_2}\, e^{-\frac{x^2}{2\sigma_2^2}}\right]$$

On peut établir facilement les relations :

$$p_1 + p_2 = 1 \qquad [1]$$

$$\mathrm{E}[|\mathrm{X}|] = \sqrt{\frac{2}{\pi}}\,[p_1\sigma_1 + p_2\sigma_2] = \mathrm{A} \qquad [2]$$

$$\mathrm{E}[\mathrm{X}^2] = p_1\sigma_1^2 + p_2\sigma_2^2 = \mathrm{B} \qquad [3]$$

$$\mathrm{E}[|\mathrm{X}|^3] = \frac{4}{\sqrt{2\pi}}\,[p_1\sigma_1^3 + p_2\sigma_2^3] = \mathrm{C} \qquad [4]$$

où A, B, C sont les moments empiriques obtenus à partir des observations.

Les quatre équations précédentes doivent nous permettre d'estimer p_1, p_2, σ_1, σ_2, à partir de $a = \mathrm{A}\sqrt{\frac{\pi}{2}}$, $b = \mathrm{B}$ et $c = \sqrt{\frac{2\pi}{16}}\,\mathrm{C}$.

$$p_1 + p_2 = 1 \qquad [5]$$

$$p_1\sigma_1 + p_2\sigma_2 = a \qquad [6]$$

$$p_1\sigma_1^2 + p_2\sigma_2^2 = b \qquad [7]$$

$$p_1\sigma_1^3 + p_2\sigma_2^3 = c \qquad [8]$$

Le produit de la relation [5] par σ_1 moins la relation [6] donne $p_2(\sigma_1 - \sigma_2) = \sigma_1 - a$.

---

[1] Le calcul de E[X] permet de se ramener au cas où $m = 0$ car $\mathrm{E}[\mathrm{X}] = mp_1 + mp_2 = m$.

En procédant de même avec les relations [6] et [7], et [7] et [8] on obtient trois relations :

$$p_2(\sigma_1 - \sigma_2) = \sigma_1 - a \qquad [9]$$

$$p_2\sigma_2(\sigma_1 - \sigma_2) = a\sigma_1 - b \qquad [10]$$

$$p_2\sigma_2^2(\sigma_1 - \sigma_2) = b\sigma_1 - c \qquad [11]$$

Le rapport de [11] sur [10] et de [10] sur [9] donne :

$$\sigma_2 = \frac{b\sigma_1 - c}{a\sigma_1 - b} \qquad [12]$$

$$\sigma_2 = \frac{a\sigma_1 - b}{\sigma_1 - a} \qquad [13]$$

L'égalisation de [12] et [13] donne l'expression de σ_1 :

$$(a\sigma_1 - b)^2 = (b\sigma_1 - c)\ (\sigma_1 - a)$$

$$\sigma_1^2(a^2 - b) - (ab - c)\,\sigma_1 + b^2 - ac = 0 \qquad [14]$$

On obtiendrait la même équation en σ_2 en multipliant [5], [6] et [7] par σ_2 au lieu de σ_1.
On a donc :

$$\sigma = \frac{ab - c \pm \sqrt{(ab-c)^2 - 4(a^2-b)(b^2-ac)}}{2(a^2-b)}$$

Les racines de [14] donnent les valeurs de σ_1 et σ_2.
Ces racines existent et sont réelles car :

$$(ab-c)^2 - 4(a^2-b)(b^2-ac) > 0$$

En effet :

$$b^2 - ac = p_1p_2\sigma_1\sigma_2(2\sigma_1\sigma_2 - \sigma_1^2 - \sigma_2^2) = -p_1p_2\sigma_1\sigma_2(\sigma_1-\sigma_2)^2 < 0$$

$$a^2 - b = (p_1^2 - p_1)\sigma_1^2 + (p_2^2 - p_2)\sigma_2^2 + 2p_1p_2\sigma_1\sigma_2 = p_1p_2(\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2) > 0$$

Donc :

$$-4(a^2-b)(b^2-ac) > 0$$

Ceci est vrai en théorie, mais avec les valeurs de a, b, c, résultant des calculs sur une distribution empirique, il peut se faire que le radical soit négatif et σ_1 et σ_2 imaginaires.

Les expressions de p_1 et p_2 s'obtiennent aisément en reportant les valeurs de σ_1 et σ_2 dans [9], puis [5]. Les résultats fournis par ces équations nous semblent cependant exiger un grand nombre d'observations, car la dispersion existant sur a, b et c entraîne une dispersion très grande sur les estimations de σ_1, σ_2, p_1 et p_2 (9) (28) (21).

C'est pourquoi l'on peut se demander si l'on n'a pas intérêt à estimer ces paramètres à partir du graphique de Henri en combinant dans diverses proportions deux lois normales d'écarts types différents et en se rapprochant par approximations successives de la loi empirique observée.

Nous donnons, figure 6, un exemple de combinaison de deux lois normales sur le graphique de Henri. Comme la droite de Henri théorique de chaque loi indique pour une abscisse t :

$$G(t) = \int_t^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx$$

on voit que, connaissait $G_1(t)$ et $G_2(t)$ ainsi que p, on peut calculer facilement la probabilité $G(t)$ pour que la variable aléatoire issue du mélange soit supérieure à t :

$$G(t) = p\,G_1(t) + (1-p)\,G_2(t)$$

Le graphique théorique de la figure 6 montre trois alignements joints par des lignes un peu incurvées. Il convient du reste de remarquer que la courbe $G(t)$ n'est pas tout à fait rectiligne sur le graphique de Henri, mais il s'en faut de peu.

D'autre part, on peut faire à propos des $k(t)$ les mêmes remarques que pour les $V(t)$. Les $k(t)$

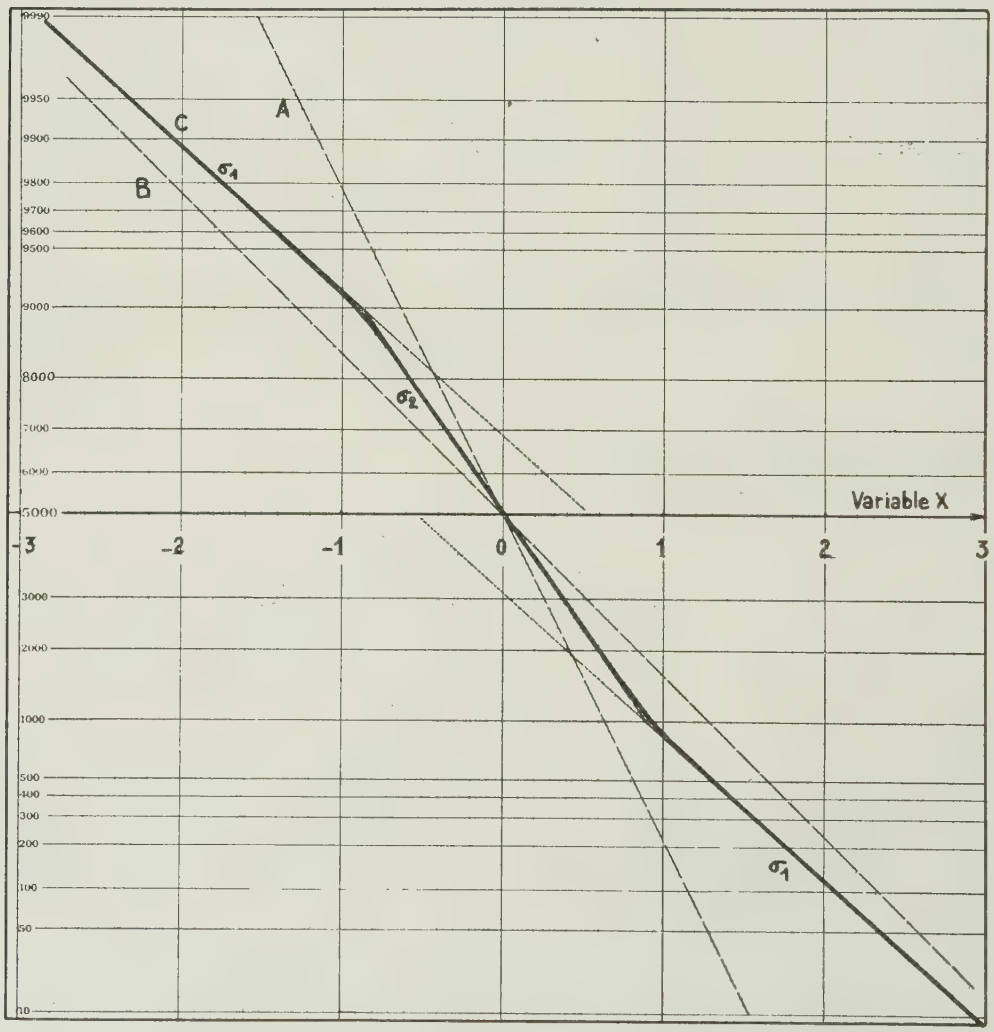


FIG. 6. — Composition à égalité
de deux lois normales de moyennes nulles et d'écarts types égaux à 1 et 1/2.

— A — Population $m = 0$, $\sigma = 1/2$.
— B — Population $m = 0$, $\sigma = 1$.
— C — Population résultante :

$m = 0$ { $\sigma_1 = 1{,}1$. $\sigma_2 = 0{,}7$.

calculés sur des V (t) moyennés sont moins dispersés que les $k\,(t)$ provenant des V (t) bruts. Enfin la dispersion s'accroît quand l'échantillonnage est moins serré.

Nous avons effectué des ajustements à la loi normale des coefficients de réflexion secteur par secteur. Les ajustements présentent les mêmes caractéristiques que plus haut, mais on remarque en outre une variation de la dispersion du simple au triple selon le secteur, ce qui dénote un certain manque de stationnarité des $k\,(t)$.

Espacements entre $|k| >$ seuil.

Comme pour les vitesses, on peut chercher la loi régissant les intervalles entre les coefficients $k\,(t)$ successifs dont la valeur absolue est supérieure à un seuil donné, c'est-à-dire la loi de répartition des couches réfléchissantes notables dans le sol. Nous avons réalisé des ajustements des espacements à des lois exponentielles (fig. 7 et (2)). Les ajustements graphiques ne sont pas tous excellents. Des tests x^2 de Pearson ont été effectués et ont conclu dans l'ensemble à l'acceptation de l'hypothèse exponentielle [Prob. (intervalle $> x$) $= e^{-\frac{x}{d}}$ où d est l'intervalle moyen entre deux coefficients successifs tels que $|k(t)| >$ C].

Autocorrélation des $k\,(t)$.

La plupart des études statistiques en sismique supposent que les coefficients de réflexion sont discrets et répartis à la Poisson, ce qui suppose leur autocorrélation nulle ailleurs qu'à l'origine. On peut observer sur la figure 2 que les $k\,(t)$ ne sont pas discrets mais qu'on peut en tracer une courbe continue. En fait, un grand nombre de coefficients sont faibles et on a vu que les coefficients importants ($|k| >$ seuil) étaient répartis à la Poisson. Nous avons calculé leur autocorrélation dans un cas particulier. La figure 8 montre l'autocorrélation des $k\,(t)$ de la figure 2. On peut vérifier que l'autocorrélation fluctue pratiquement autour de 0 avec des écarts de $\pm$ 0,1. Ce fait coïncide bien avec la constatation de la répartition poissonienne des $|k\,(t)| >$ C.

Remarque : intervalles de confiance dans les ajustements graphiques.

Pour juger la signification des divers ajustements graphiques précédents (graphique de Henri et ajustements exponentiels) on peut utiliser la méthode du chenal de confiance. La figure 9 indique les limites entre lesquelles peuvent évoluer les pourcentages observés dans un exemple d'ajustement par la droite de Henri. Les limites du chenal n'ont que 5 % de chances d'être dépassées si la population testée obéit à la loi normale. Les limites dépendent du nombre des observations de la population. Elles se resserrent quand la population s'accroît.

La théorie des chenaux de confiance est simple. Si la proportion des observations supérieures à une valeur x est en théorie p et que l'on dispose de N observations, on sait que la loi du nombre n_i des observations supérieures à x est fournie par la loi binomiale :

$$\text{Prob.}\,(n_i) = p^{n_i}(1-p)^{N-n_i}C_N^{n_i} \qquad 0 \leqslant n_i \leqslant N$$

On peut ainsi calculer pour N et p fixés les nombres n_i' et n_i'' tels que :

$$\text{Prob.}\,(n_i > n_i') = 2{,}5\ \%$$
$$\text{Prob.}\,(n_i > n_i'') = 97{,}5\ \%$$

Ce sont les courbes $\frac{n_i'}{N}$ et $\frac{n_i''}{N}$ qui sont portées sur la figure 9 dans le cas d'une loi normale de variance σ, de moyenne nulle, pour N = 10, 50, 100, 500.

Thèse J. AGARD. II.

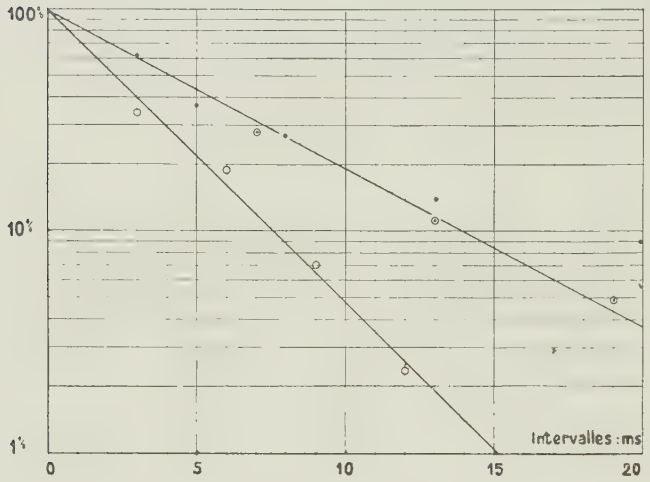


FIG. 7. — Ajustement exponentiel.
Intervalles entre coeff. de réflexion.

$|k| >$ seuil

● 74 observations $|k| > 0{,}045$.
⊙ 82 observations $|k| > 0{,}040$.
○ 127 observations $|k| > 0{,}025$.

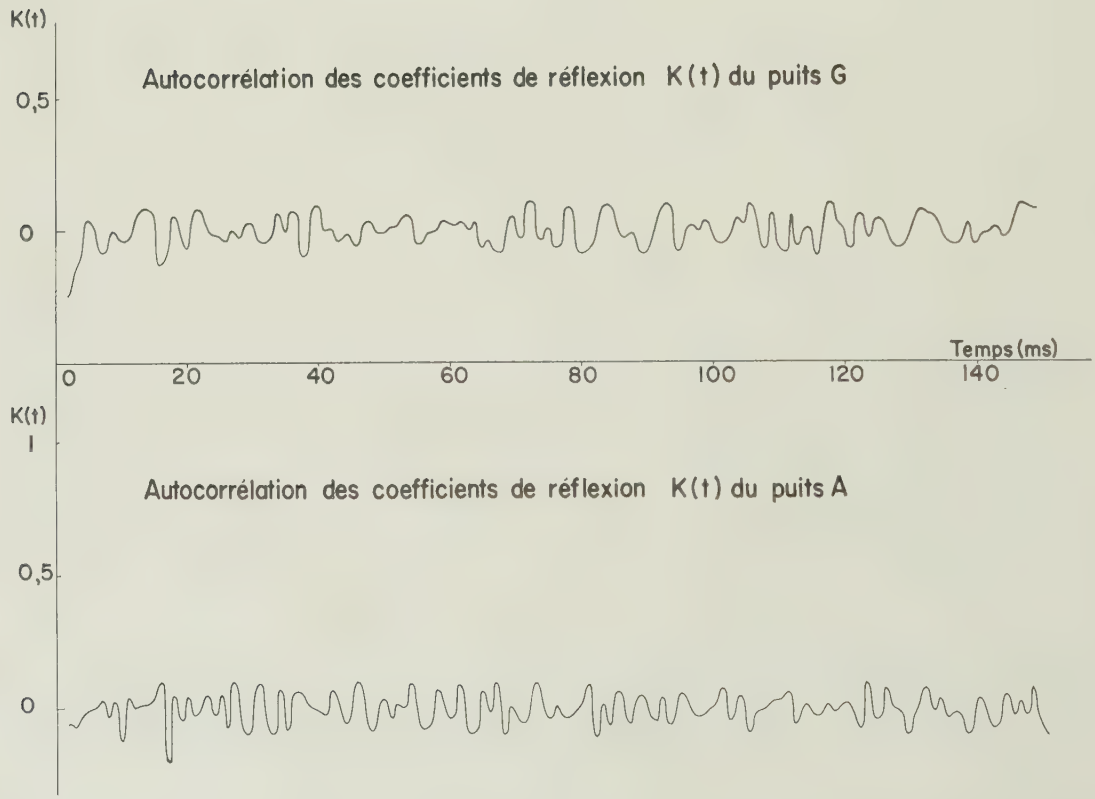


FIG. 8.

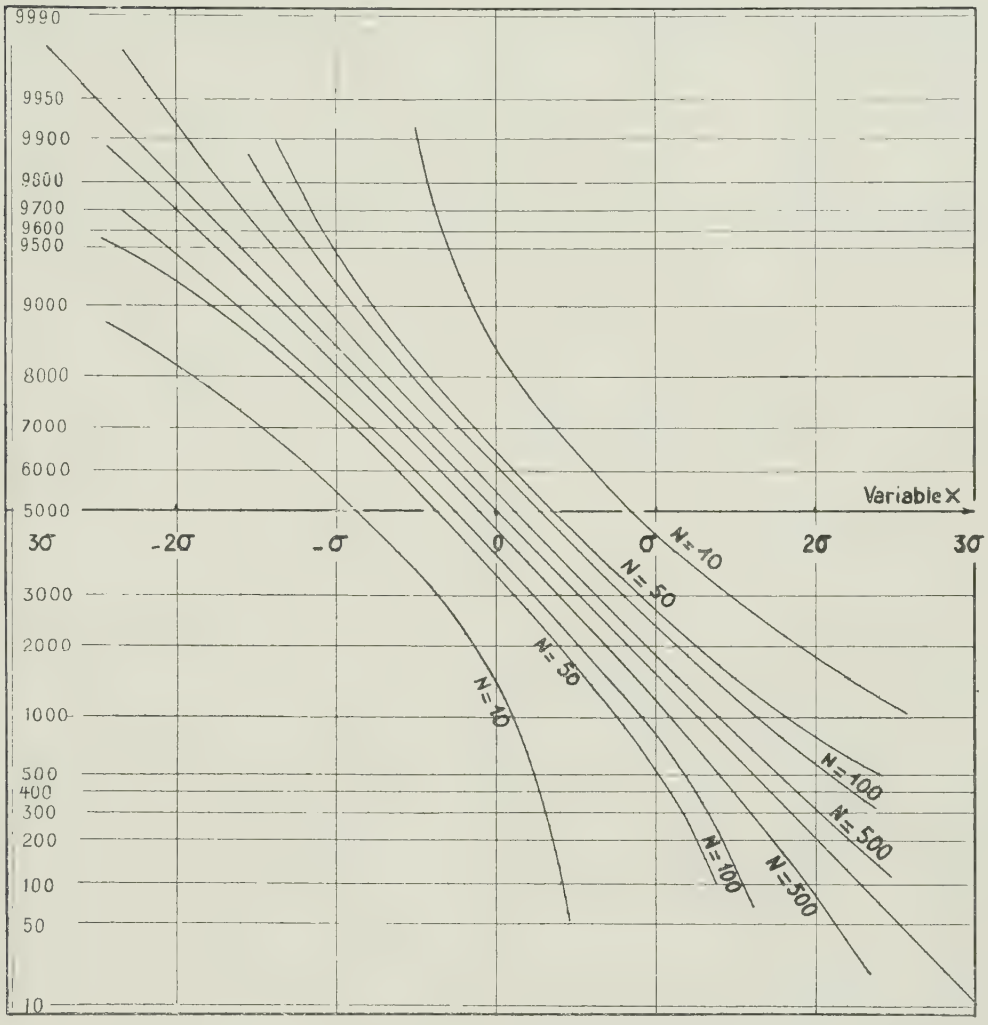


FIG. 9. — Intervalles de confiance à 95 % en fonction du nombre des observations.

La construction des chenaux de confiance est d'une signification moins nette que la valeur du χ^2, mais elle est rapide et donne une bonne indication sur la validité d'un ajustement graphique [24].

b. Étude théorique de la répartition des k (t).

Les ajustements effectués sur les coefficients de réflexion s'expliquent en partie par le schéma théorique suivant où l'on examine quelques fonctions de deux variables normales x et y corrélées.

Nous étudions les densités de probabilité de fonctions de deux variables normales corrélées x et y. Les fonctions analysées ici sont :

La somme $x + y$;

La différence $x - y$;

Le couple $x + y$, $x - y$;

Le produit xy ;

Le rapport $\frac{x}{y}$;

Le rapport $\frac{x-y}{x+y}$.

1) Ces quelques fonctions se rencontrent en particulier dans le cas d'une série temporelle $x(t)$, $y(t) = x(t+h)$ où $x(t)$ peut être considéré comme une variable normale à autocorrélation non nulle. Par exemple, si $x(t)$ est la vitesse de propagation d'un milieu, on peut s'intéresser à la loi de répartition des coefficients de réflexion : $\frac{x(t)-x(t+h)}{x(t)+x(t+h)}$, sachant que $x(t)$ est une loi normale dont l'autocorrélation est connue. Ces problèmes se rencontrent dans le cas du radar monopulse (14).

2) *Loi de $x+y$.*

La densité de probabilité du couple (x, y) est :

$$f(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{(x-m)^2+(y-m)^2-2r(x-m)(y-m)}{2\sigma^2(1-r^2)}}$$

On suppose dans tout ce qui suit que x et y ont même moyenne m, même variance σ^2 et une corrélation r.

Pour déterminer la densité de probabilité d'une fonction $u = g(x, y)$, on fait en général un changement de variable pour passer du couple (x, y) au couple (u, x) ou (u, y) ou (u, v) avec $v = h(x, y)$. On intègre ensuite par rapport à la seconde variable et on obtient la densité de probabilité de la variable u.

Dans le cas de la somme $u = x + y$, on retrouve le résultat bien connu : u est une variable normale.

Le couple (u, x) a pour densité de probabilité :

$$\frac{1}{2\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{(x-m)^2+(u-x-m)^2-2r(x-m)(u-x-m)}{2\sigma^2(1-r^2)}}\, du\, dx$$

$$= \frac{1}{2\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{2(1-r)\left(x-\frac{u}{2}\right)^2+(1-r)\left(\frac{u^2}{2}+2m^2-2mu\right)}{2\sigma^2(1-r^2)}}\, dx\, du$$

En intégrant cette expression par rapport à x, ce qui est possible par un nouveau changement de variable $\left(t = x - \frac{u}{2}\right)$, on obtient pour loi de u :

$$f(u) = \frac{1}{2\sigma\sqrt{(1+r)\pi}}\, e^{-\frac{(u-2m)^2}{4\sigma^2(1+r)}}$$

La variable $u = x + y$ est une variable normale de moyenne $2m$ et d'écart-type $\sigma\sqrt{2(1+r)}$. Quand $r = 0$, on retrouve le résultat classique d'addition de deux variables indépendantes.

3) *Loi de $x - y$.*

Par un calcul analogue au précédent, on obtient pour densité de probabilité de la différence $u = x - y$:

$$f(u) = \frac{1}{2\sigma\sqrt{(1-r)\pi}}\, e^{-\frac{u^2}{4\sigma^2(1-r)}}$$

La différence u est une variable normale de moyenne nulle et d'écart type $\sigma\sqrt{2(1-r)}$.

4) *Loi du couple $x - y$, $x + y$.*

Effectuons le changement de variable $\begin{cases} u = x + y \\ v = x - y \end{cases}$

La densité de probabilité du couple (u, v) est, avec le Jacobien $\dfrac{D(u,v)}{D(x,y)} = \dfrac{1}{2}$ et $x = \dfrac{u+v}{2}$, $y = \dfrac{u-v}{2}$:

$$\frac{1}{4\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{\left(\frac{u+v}{2}-m\right)^2+\left(\frac{u-v}{2}-m\right)^2-2r\left(\frac{u+v}{2}-m\right)\left(\frac{u-v}{2}-m\right)}{2\sigma^2(1-r^2)}}$$

$$= \frac{1}{4\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{(u-2m)^2(1-r)+v^2(1+r)}{4\sigma^2(1-r^2)}}$$

On retrouve les résultats exposés aux paragraphes 2 et 3 : loi de $x + y$ et loi de $x - y$. De plus, on remarque que u et v sont deux variables aléatoires indépendantes, car leurs densités se séparent dans la densité du couple (u, v). Si les valeurs moyennes ou les écarts types des deux variables x et y ne se trouvaient pas égaux, les termes rectangles en uv ne disparaîtraient pas et u et v seraient deux variables normales corrélées.

5) *Loi du produit xy.*

La densité de probabilité du couple $xy = u$ et de x est, dans le cas de moyenne nulle de x et y :

$$\frac{1}{2\pi\sigma^2\sqrt{1-r^2}\,|x|}\, e^{-(x^2+\frac{u^2}{x^2}-2ru)\frac{1}{2\sigma^2(1-r^2)}} \quad \left(\frac{D(u,v)}{D(x,y)} = \left|\frac{1}{x}\right|\right)$$

La densité de probabilité de u s'obtient par intégration sur x :

$$\int_{-\infty}^{\infty} \frac{e^{-(x^2+\frac{u^2}{x^2}-2ru)\frac{1}{2\sigma^2(1-r^2)}}}{2\pi\sigma^2\sqrt{1-r^2}\,|x|}\, dx = \frac{e^{-\frac{ru}{\sigma^2(1-r^2)}}}{\pi\sigma^2\sqrt{1-r^2}}\, k_0\left(\frac{|u|}{1-r^2}\right)$$

où k_0 est la fonction de Bessel Hankel d'ordre 0.

Si les variables x et y sont indépendantes ($r = 0$), on a :

$$f(u) = \frac{1}{\pi\sigma^2}\, k_0(|u|)$$

Si la corrélation $r = 1$, la loi de u est une loi du type III de Pearson : $f(u) = \dfrac{e^{-\frac{|u|}{2}}}{\sqrt{2\pi}\,\sigma|u|}$

En utilisant les tables de Bessel, on peut calculer $f(u)$ pour diverses valeurs de r. MORLAT a construit quelques courbes de ce type pour $r = 0$; 0,5 ; 0,9 ; 1.

6) *Loi du rapport* $\dfrac{x}{y}$.

La densité de probabilité du couple $\dfrac{x}{y} = u$ et de y est obtenue à partir de la densité du couple (x, y) du paragraphe 2 : loi de $x + y$, avec un Jacobien $\dfrac{D(u,y)}{D(x,y)} = |y|$.

$$\frac{1}{2\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{[(uy-m)^2+(y-m)^2-2r(uy-m)(y-m)]}{2\sigma^2(1-r^2)}}\,|y|\,dy\,du$$

$$= \frac{1}{2\pi\sigma^2\sqrt{1-r^2}}\, e^{-\left[\left(y\sqrt{u^2-2ru+1}-\frac{m(1-r)(1+u)}{\sqrt{u^2-2ur+1}}\right)^2+\frac{(1-r^2)(u-1)^2m^2}{u^2-2ru+1}\right]\frac{1}{2\sigma^2(1-r^2)}}\,|y|\,dy\,du$$

$$f(u) = \frac{e^{-\frac{(u-1)^2m^2}{(u^2-2ru+1)2\sigma^2}}}{2\pi\sigma^2\sqrt{1-r^2}} \int_{-\infty}^{\infty} e^{-\left(y\sqrt{u^2-2ru+1}-\frac{m(1-r)(u+1)}{\sqrt{u^2-2ur+1}}\right)^2\frac{1}{2\sigma^2(1-r^2)}}\,|y|\,dy$$

Le calcul de $\int_{-\infty}^{\infty} e^{-(y-a)^2 b}\,|y|\,dy$ donne :

$$\int_0^\infty e^{-(y-a)^2 b}\,y\,dy + \int_0^\infty e^{-(y+a)^2 b}\,y\,dy = \int_0^\infty e^{-(y-a)^2 b}\,(y-a+a)\,dy + \int_0^\infty e^{-(y+a)^2 b}\,(y+a-a)\,dy$$

$$\int_{a^2}^\infty e^{-Xb}\,\frac{dX}{2} + a\int_{-a}^\infty e^{-Y^2 b}\,dY + \int_{a^2}^\infty e^{-X'b}\,\frac{dX'}{2} - a\int_a^\infty e^{-Y'^2 b}\,dY'$$

avec $X = (y-a)^2$; $Y = y-a$; $X' = (y+a)^2$; $Y' = y+a$

$$= \frac{e^{-ba^2}}{2b} + \frac{a\sqrt{2\pi}}{\sqrt{2b}}\,G(-a) + \frac{e^{-ba^2}}{2b} - \frac{a\sqrt{2\pi}}{\sqrt{2b}}\,G(a) = \frac{e^{-ba^2}}{b} + a\sqrt{\frac{\pi}{b}}\,[G(-a)-G(a)]$$

$$\left(\text{où } G(a) = \frac{\sqrt{2b}}{\sqrt{2\pi}}\int_a^\infty e^{-\frac{x^2 2b}{2}}\,dx \text{ est la loi de Gauss de variance } \frac{1}{b} \text{ et } A = e^{-\frac{m^2(1-r)(u+1)^2}{2\sigma^2(1+r)(u^2-2ru+1)}}\right)$$

$$f(u) = \frac{e^{-\frac{(u-1)^2m^2}{2\sigma^2(u^2-2ur+1)}}}{2\pi\sigma^2\sqrt{1-r^2}}\left[\frac{2\sigma^2(1-r^2)A}{u^2-2ur+1} + \frac{m(1-r)(u+1)}{u^2-2ur+1}\sqrt{\frac{2\pi\sigma^2(1-r^2)}{u^2-2ur+1}}\,\Pi\right]$$

avec :

$$\Pi = \int_{-\frac{m(1-r)(u+1)}{u^2-2ur+1}}^{\frac{m(1-r)(u+1)}{u^2-2ur+1}} e^{-\frac{x^2}{2}\left[\frac{u^2-2ur+1}{\sigma^2(1-r^2)}\right]} \frac{dx}{\sqrt{2\pi}\sqrt{\frac{\sigma^2(1-r^2)}{u^2-2ur+1}}}$$

Lorsque $m = 0$, on trouve la loi de Cauchy :

$$f(u) = \frac{\sqrt{1-r^2}}{(u^2-2ur+1)\pi}$$

la variable réduite est $\frac{u-r}{\sqrt{1-r^2}}$.

Si $r = 1$, on a toujours $u = 1$.

Si m devient très grand par rapport à σ, Π et A sont est pratiquement égaux à 1 et u est distribué pratiquement comme une loi normale de moyenne 1 et d'écart type $\frac{\sigma\sqrt{2}}{m\sqrt{1-r}}$.

7) *Loi du rapport* $\frac{x-y}{x+y}$.

On recherche la densité de probabilité du couple $\left(u = \frac{x-y}{x+y}, y\right)$ à partir de celle du couple (x, y). On a :

$$\frac{1}{2\pi\sigma^2\sqrt{1-r^2}} e^{-\left[\left(y\frac{1+u}{1-u}-m\right)^2+(y-m)^2-2r\left(y\frac{1+u}{1-u}-m\right)(y-m)\right]\frac{1}{2\sigma^2(1-r^2)}} \frac{2|y|}{(1-u)^2}\,dy\,du$$

$\left(\text{le Jacobien est égal à } \frac{2|y|}{(1-u)^2}\right)$.

On peut transformer cette expression en :

$$\frac{1}{2\pi\sigma^2\sqrt{1-r^2}} e^{\frac{-\frac{m^2u^2}{\sigma^2(1+u^2-r(1-u^2))}}{(1-u^2)}} \int_{-\infty}^{\infty} e^{-\left(y-\frac{m(1-r)(1-u)}{u^2+1-r(1-u^2)}\right)^2 \frac{u^2+1-r(1-u^2)}{\sigma^2(1-r^2)(1-u)^2}} 2|y|dy\,du$$

Il reste à effectuer un calcul analogue à celui du paragraphe 6 : loi du rapport $\frac{x}{y}$, et on trouve pour densité $f(u)$:

$$f(u) = \frac{1}{\pi\sigma^2\sqrt{1-r^2}} e^{\frac{-\frac{m^2u^2}{1+u^2-r(1-u^2)}}{(1-u)^2}} \left[\frac{\sigma^2(1-r^2)(1-u)^2}{u^2+1-r(1-u^2)} A' + \frac{\Pi m(1-r)(1-u)}{u^2+1-r(1-u^2)}\sqrt{\frac{\pi\sigma^2(1-r)(1-u^2)}{u^2+1-r(1-u^2)}}\right]$$

$$\Pi = \int_{-\frac{m(1-r)(1-u)}{u^2+1-r(1-u^2)}}^{\frac{m(1-r)(1-u)}{u^2+1-r(1-u^2)}} \frac{e^{-\frac{x^2}{2}\left[\frac{(u^2+1-r(1-u^2))^2}{\sigma^2(1-r^2)(1-u)^2}\right]}}{\sqrt{2\pi}\sqrt{\frac{\sigma^2(1-r^2)(1-u)^2}{2(u^2+1-r(1-u))^2}}}, \quad A' = e^{-\frac{m^2(1-r)}{\sigma^2(1+r)[u^2+1-r(1-u^2)]}}$$

Si $m = 0$, on retrouve la loi de Cauchy :

$$f(u) = \frac{1 + r}{\pi\left(u^2 \frac{(1 + r)}{1 - r} + 1\right)}$$

La variable réduite est $u\sqrt{\frac{1 + r}{1 - r}}$.

Si $r = 1$, la variable $u = 0$.

Si m est très grand devant σ, Π et A' sont pratiquement égaux à 1 et u est pratiquement distribué comme une loi normale de moyenne 0 et d'écart type $\frac{\sigma\sqrt{1 - r}}{\sqrt{2m}}$.

Vérifications sur un exemple synthétique.

Pour vérifier les considérations précédentes sur un exemple simple, nous avons constitué un échantillon de vitesses normales V (t) de moyenne 3 500 m/s et d'écart type $\sigma_V = 950$ m/s. Nous avons tiré 300 vitesses au hasard suivant cette loi (avec une corrélation nulle entre les V (t). La figure 10

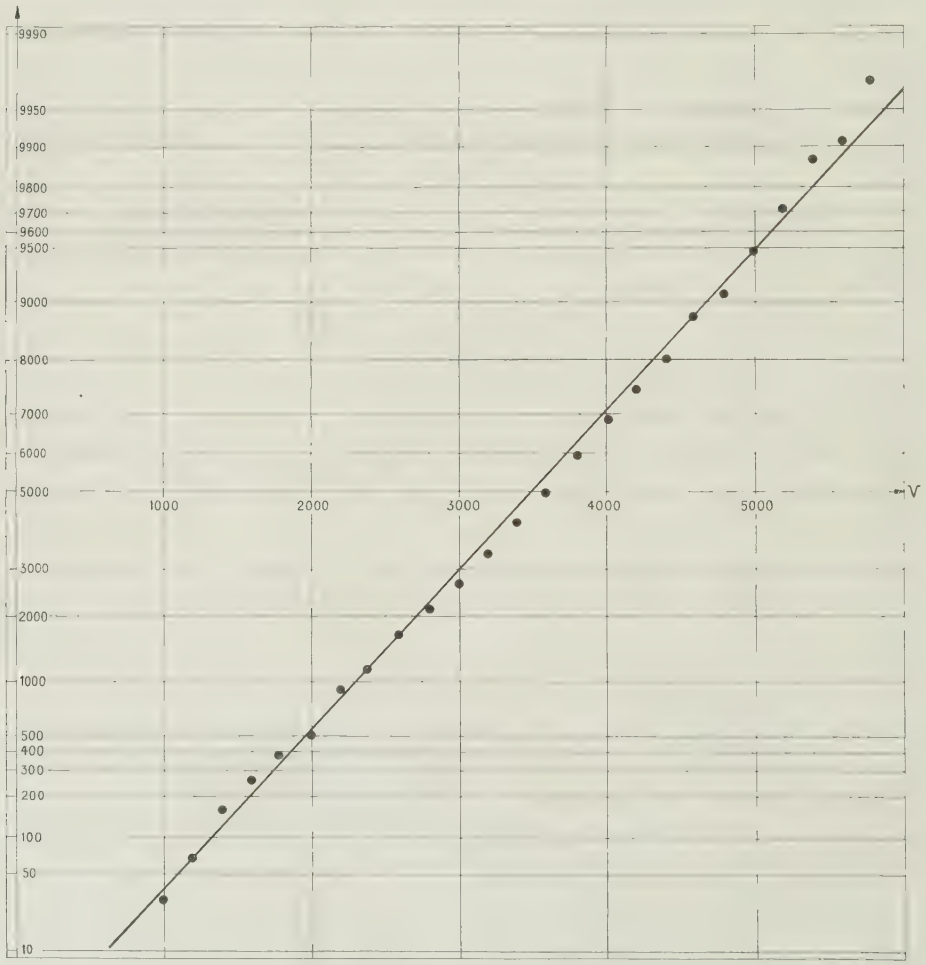


Fig. 10. — Échantillon de vitesses synthétiques suivant la loi normale.
$m = 3\,500$. $\sigma = 900$. $n = 300$.

montre la droite de Henri théorique et les points représentatifs de l'échantillonnage. A partir de cet échantillonnage, on a construit 299 valeurs k par la formule :

$$k_j = \frac{V_j - V_{j+1}}{V_j + V_{j+1}}$$

On constate que les k_j peuvent être assimilés à une loi normale sur le graphique de la figure 11.

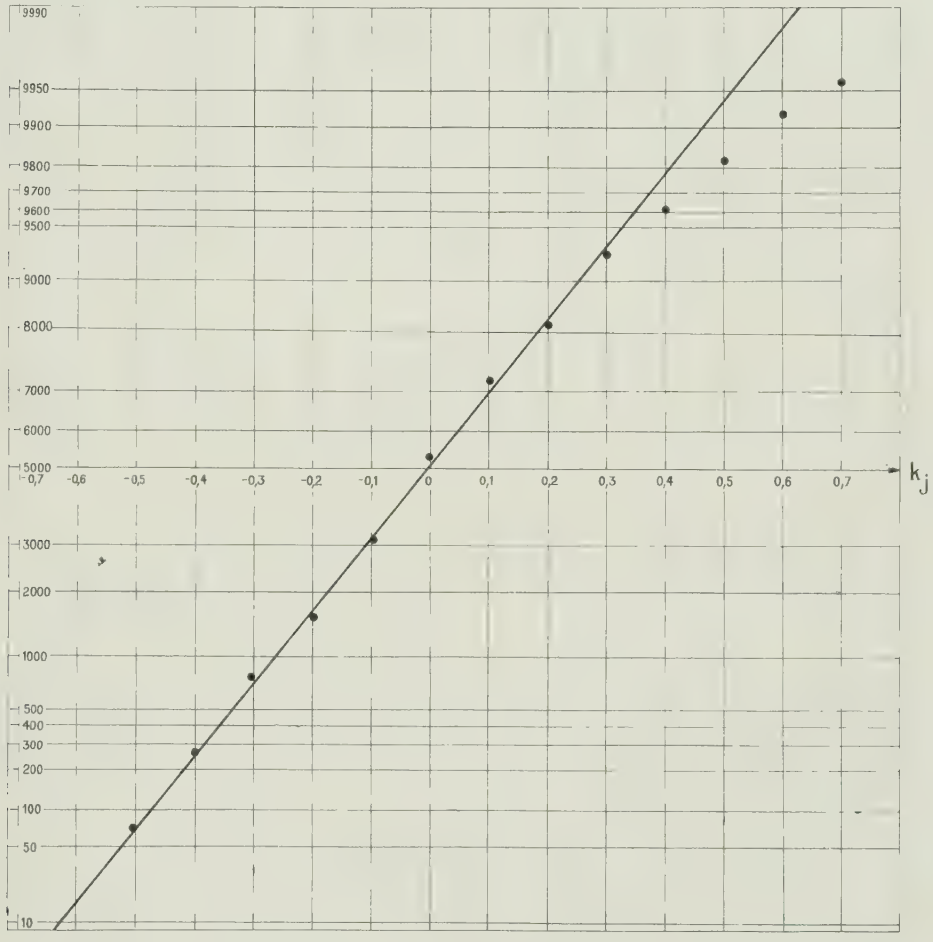


FIG. 11. — Ajustement à la loi normale des coefficients de réflexion synthétiques.

$$k_j = \frac{V_j - V_j + 1}{V_j + V_j + 1}$$

$m = 0.$ $\sigma = 0{,}21.$ $n = 299.$

La moyenne des k (t) dans les deux cas est nulle et leur écart type est de 0,20 et 0,22 alors que l'on devrait avoir $\sigma = \frac{\sqrt{2} \cdot \sigma_V}{2m} = \frac{\sqrt{2} \cdot 950}{2 \cdot 3\,500} = 0{,}192$ d'après les formules théoriques précédentes. L'accord est très satisfaisant.

Enfin, en translatant toutes les vitesses de l'échantillon précédent pour leur donner une vitesse moyenne nulle et en calculant les $k_j = \frac{V_j - V_{j+1}}{V_j + V_{j+1}}$ on obtient un ajustement excellent à une loi de Cauchy. Les formules théoriques sont donc bien vérifiées.

3. Étude de l'impulsion $s(t)$.

L'impulsion $s(t)$ engendrée par l'explosion est l'une des inconnues des sismogrammes. Théoriquement, cette impulsion est à l'origine un Dirac. En fait ce Dirac s'étale très vite et le signal $s(t)$ enregistré par les sismographes a une longueur variant de 50 à 100 ms environ. La longueur du signal augmente un peu au cours de la propagation.

La forme analytique de l'impulsion $s(t)$ n'est pas connue. Certains proposent la forme $e^{-\alpha t}\sin\sqrt{t}$. D'autres comme HORTON utilisent un développement en série de Gram-Charlier (7) :

$$s(t) = \sum_n b_n \varphi^{(n)}\left(\frac{t}{T}\right) \qquad \text{avec} \qquad \varphi^{(n)}(x) = \frac{d^n}{dx^n}\left(\frac{e^{-x^2/2}}{\sqrt{2\pi}}\right)$$

RICKER propose la forme normée :

$$s(t) = \frac{\sqrt{\pi}}{m}\left(1 - \frac{2t^2}{m^2}\right) e^{\frac{t^2}{m^2}}$$

Dans certains cas, une simple période de sinusoïde peut suffire. Enfin, il arrive qu'une ou deux réflexions très importantes du sismogramme nous donnent l'image de $s(t)$.

4. Étude des bruits.

Nous avons exposé plus haut les trois sortes de bruit rencontrés en sismique. Les bruits induits par une ligne à haute tension voisine du tir, les ondes de surface entretenues (*groundroll*), ou les oscillations obtenues avec des tirs dans l'eau ont une structure périodique qu'on peut représenter par une fonction $g(t)$ inconnue a priori.

D'autres bruits correspondent à des bruits divers en surface ou à l'enregistrement et ils n'ont aucune structure périodique. Nous les représentons par $\varepsilon(t)$. On examinera au chapitre III les résultats statistiques d'une expérimentation propre à éliminer ces bruits quasi périodiques ou non.

5. Étude des amplitudes des sismogrammes.

On peut a priori supposer que les amplitudes d'un sismogramme :

$$y(t) = \int_{t-\infty}^{t} s(t-\tau)\, k\left(\frac{\tau}{2}\right) d\tau + g(t) + \varepsilon(t)$$

suivent une loi normale (si l'on néglige l'absorption ou si l'on a utilisé un contrôle de gain). En effet, les réflexions peuvent être comparées à un bruit aléatoire :

$$I(t) = \sum_{-\infty}^{\infty} a_k F(t - t_k)$$

où les t_k sont les instants de départ d'impulsions à la Poisson de densité ν et les a_k des variables aléatoires de même loi. RICE (24) montre que $I(t)$ tend vers une loi normale quand $\nu \to \infty$. Les $k\left(\frac{t}{2}\right)$ remplissant les conditions fixées pour les a_k, le premier terme de $y(t)$ suit une loi normale.

Le bruit $\varepsilon(t)$ peut être supposé un bruit blanc. Si $g(t)$ n'a pas une amplitude prédominante, $y(t)$ devrait donc suivre approximativement une loi normale. Nous avons vérifié la normalité dans le cas de sismogrammes synthétiques (chapitre IX) pour lesquels $g(t)$ et $\varepsilon(t)$ n'existent pas. La figure 12 montre une droite de Henri bien ajustée. La loi normale est bien la loi des amplitudes. On peut remarquer également que la densité de l'échantillonnage ne change rien à la loi.

Nous avons vérifié la même propriété sur des sismogrammes réels (2) où un bruit organisé $g(t)$ était peu probable.

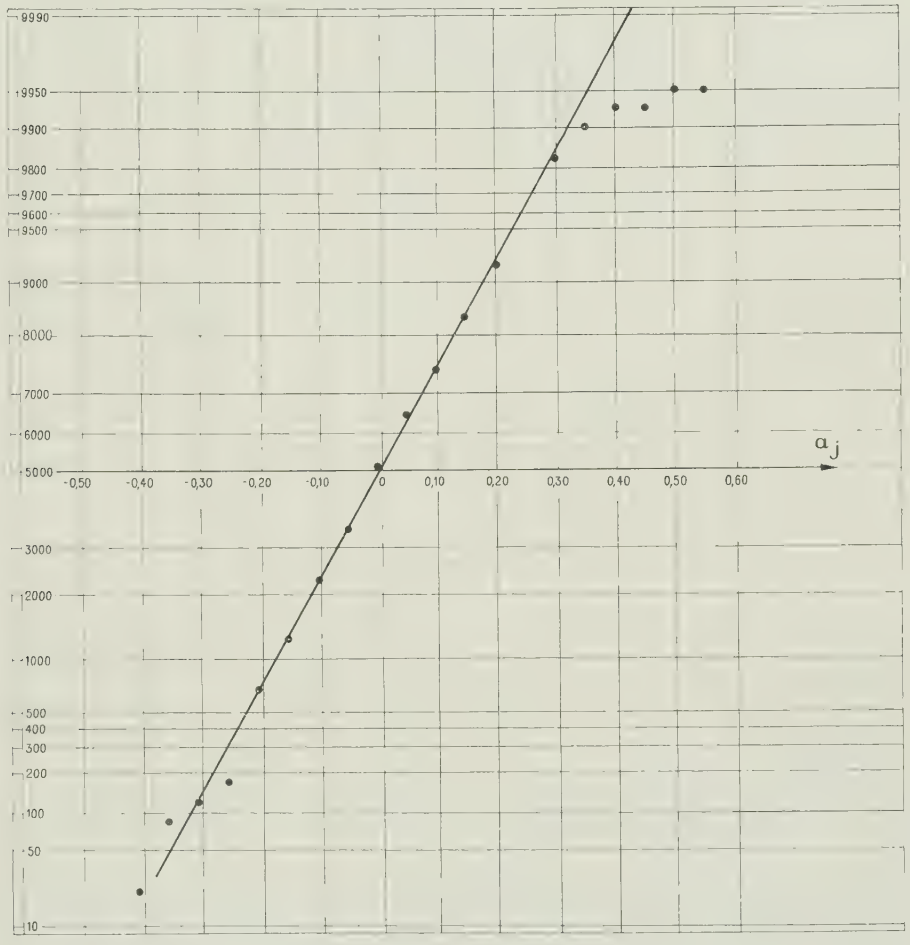


FIG. 12. — Ajustement à la loi normale des amplitudes du sismogramme d'Erg Chech 171.

$m = 0.$ $\sigma = 0{,}14.$ $n = 600.$

Effet de $g(t)$.

Dans le cas où un bruit organisé important $g(t)$ se manifeste, la loi normale ne régit plus les amplitudes. Pour tester cette influence sur un exemple simple, nous avons ajouté une sinusoïde aux amplitudes d'un sismogramme synthétique :

$$y'(t) = \int_{t-\mathfrak{S}}^{t} k\left(\frac{\tau}{2}\right) s(t-\tau)\, d\tau + a \cos \omega (t + \varphi)$$

Si nous cherchons la loi de $y(t)$ en un point t fixé, en faisant abstraction des valeurs prises par $y(t)$ avant t, on aura pour densité de répartition $f(y)$:

$$f(y) = \int_{-a}^{a} \frac{e^{-\frac{(y-u)^2}{2\sigma^2}}}{\pi\sqrt{2\pi}\sigma\sqrt{a^2-u^2}}\,du \qquad (u = a\cos\omega(t+\varphi))$$

En effet, on a vu plus haut que $x = \int_{t-\theta}^{t} k\left(\frac{\tau}{2}\right) s(t-\tau)\,d\tau$ était réparti comme une variable normale de moyenne nulle. Or, la densité de répartition de $u = a\cos\omega(t+\varphi)$ en un point t fixé quand la phase a une répartition uniforme entre o et $\frac{\pi}{\omega}$ est :

$$\psi(u) = \frac{1}{\pi\sqrt{a^2-u^2}}$$

On voit que la fréquence ω n'intervient pas dans $f(y)$. L'amplitude a joué par contre un grand rôle. Si a est petit, $f(y)$ est approximativement une densité normale. Si a devient très grand par rapport à σ, $f(y)$ tend vers la densité de u.

On a :

$$E(Y) = E(X) + E(U) = o$$

$$E(Y^2) = E(X^2) + E(U^2) = \sigma^2 + \frac{a^2}{2}$$

en supposant X et U sans corrélation. Sinon, on a :

$$E(Y^2) = E(X^2) + E(U^2) + 2\rho\sqrt{E(X^2)\cdot E(U^2)}$$

où ρ est la corrélation normée de U et de X dans laquelle a n'intervient pas.

On voit que, pour que $E(Y^2)$ soit minimal, il faut que $a + \rho\sigma\sqrt{2} = o$, soit $a = -\rho\sigma\sqrt{2}$.

Si la sinusoïde $u = \sin\omega(t+\varphi)$ d'amplitude 1 a une corrélation $\rho > o$ avec X, il faudra donner pour amplitude $\rho\sigma\sqrt{2}$ à la sinusoïde et la déphaser de $\frac{\pi}{\omega}$ pour mettre u en corrélation négative avec X et minimiser $E(Y^2) = \sigma^2 - \rho^2\sigma^2$.

Si $\rho < o$, il suffira de prendre pour amplitude de la sinusoïde $\rho\sigma\sqrt{2}$ pour que $E(Y^2) = \sigma^2 - \rho^2\sigma^2$ soit minimal (1).

Si $\rho = o$, on peut montrer que tous les moments impairs de y sont nuls et que le moment d'ordre $2n$ est égal à :

$$E(Y^{2n}) = E(X^{2n}) + \ldots + C_{2n}^{2k} E[X^{2n-2k}]\,E[U^{2k}] + \ldots + E(U^{2n})$$

$$= (2n-1)(2n-3)\ldots 3\,\sigma^{2n} + \ldots + \frac{(2n)!\,(2n-2k-1)\ldots 3}{(2n-2k)!\,(2k)!}\,\frac{\sigma^{2n-k}\ldots(2k-1)\ldots 3}{2k(2k-2)\ldots 2}\,a^{2k}$$

$$+ \ldots + \frac{(2n-1)(2n-3)\ldots 3}{2n(2n-2)\ldots 4\cdot 2}\,a^{2n}$$

Nous avons ajouté des sinusoïdes d'amplitudes et de fréquences différentes à un sismogramme

(1) Généralement ρ est théoriquement égal à o. Les variations de ρ lorsque la phase φ varie sont assez faibles.

synthétique. Les résultats de ces ajustements, condensés dans le tableau I, montrent bien que la fréquence n'intervient pas, que le σ_y obéit bien à la loi $\sigma_y = \sqrt{\sigma^2 + \frac{a^2}{2}}$ et que, si a devient un peu grand, la loi normale s'ajuste mal à la loi réelle de y.

Tableau I.

EXPÉRIENCES	σ_x	a	FRÉQ. (Hz)	σ_u	$\sigma_{u\,théor}$	x^2 (%)	n
1	0,17	0,30	10	0,26	0,27	$<$ 1	267
2	—	—	20	0,28	—	$<$ 1	279
3	—	0,10	20	0,18	0,18	$>$ 10	283
4	—	—	40	0,175	—	$>$ 10	275
5	—	0,30	31	0,275	0,27		285
6	—	—	31	0,25	—		285

$\sigma_x = 0{,}17$ est l'écart type des amplitudes du sismogramme brut ;
a = amplitude de la sinusoïde ajoutée ;
σ_u = écart-type du sismogramme résultant relevé sur le graphique de Henri ;
$\sigma^2_{u\,théor} = \sigma^2_x + \frac{a^2}{2}$;
n = nombre d'observations.

Rappelons que l'écart type théorique du σ est égal à $\sqrt{\frac{\sigma^2}{2n}}$. De ce fait, on peut constater que les amplitudes des expériences 1, 2, 5 et 6 ont bien même écart type, mais elles ne peuvent avoir le même écart type [1] que les amplitudes des expériences 3 et 4. On constate que les amplitudes des expériences 1 et 2 n'obéissent pas à la loi normale d'après le test χ^2. En revanche, 5 et 6 forment des populations presque normales. Cela vient peut-être de ce que nous avons échantillonné $y\,(t)$ toutes les 4 ms. De ce fait, nous ajoutons à $x\,(t)$ une sinusoïde échantillonnée de valeurs (0, + 0,15, + 0,30, + 0,15, 0 ...) dont la densité de répartition discrète minimise le pourcentage des grandes valeurs au profit des valeurs voisines de zéro. La distribution des $u\,(t)$ se rapproche ainsi plus d'une distribution normale que d'une distribution de sinusoïde. Dans l'expérience 6 on a cherché à mettre $x\,(t)$ et $u\,(t)$ en opposition de phase sur un intervalle d'environ une seconde. On a ainsi un peu réduit la dispersion de $y\,(t)$ mais dans une très faible mesure, la corrélation entre $x\,(t)$ et $u\,(t)$ étant peu élevée.

III. ÉTUDE STATISTIQUE DES AMPLITUDES

Le but de l'analyse des sismogrammes est de restituer en particulier la courbe des coefficients de réflexion $k\,(t)$ et de leurs multiples ou tout au moins de déterminer la valeur et l'emplacement des $k\,(t)$ les plus grands correspondant à un changement de couche géologique importante.

[1] Intervalle de confiance 95 % de $\sigma = 0{,}27 \pm 2.0{,}011$.

1. Étude d'un cas simple $y(t) = s(t - t_0) + \varepsilon(t)$.

Déterminons l'origine t_0 d'un signal $s(t)$ de forme connue au milieu d'un bruit $\varepsilon(t)$ stationnaire de densité de probabilité $\varphi(\varepsilon)$. Si l'origine t_0 de $s(t)$ a pour densité de probabilité $f(\tau)$, la densité de probabilité de $y(t)$ sera, en t :

$$\psi(t, y) = \int_{t-\mathfrak{C}}^{t} \varphi[y - s(t-\tau)] f(\tau)\, d\tau + \varphi(y) \left[1 - \int_{t-\mathfrak{C}}^{t} f(\tau)\, d\tau\right]$$

a. Détermination de $s(t)$ par le maximum likehood.

Échantillonnons $y(t)$ sur son intervalle de variation en t, $(t_1, t_2 \ldots t_n)$ en supposant l'espacement entre t_i, t_{i+1} suffisant pour éliminer la corrélation entre les $\varepsilon(t)$. En l'absence de $s(t - t_0)$, la loi des n observations de $y(t)$ serait :

$$\varphi(y(t_1))\, \varphi(y(t_2)) \ldots \varphi(y(t_n))\, dy(t_1) \ldots dy(t_n)$$

Si $s(t)$ figure entre t_j et t_{j+1}, la loi des n observations est, si t_j est a priori équiprobable de t_1 à t_n :

$$\varphi(y(t_1)) \ldots \varphi(y(t_j) - s_0) \ldots \varphi(y(t_{j+1} - s_1) \ldots \varphi(y(t_n))\, dy(t_1) \ldots dy(t_n)$$

Il faut trouver la valeur t_j qui maximise la probabilité précédente. Si la densité de probabilité de $\varepsilon(t)$ est de la forme $Ke^{-|\varepsilon|A}$ ou $Ke^{-\varepsilon^2 A}$, la maximisation du produit précédent revient à minimiser :

$$H(t) = \int_0^{\mathfrak{C}} |y(t+\tau) - s(\tau)| d\tau + \int_{t_1}^{t} |y(t)| d\tau + \int_{t+\mathfrak{C}}^{t_n = t_1 + T} |y(t)| dt$$

(où T est l'intervalle de variation de $y(t)$ entre t_1 et t_n et $\mathfrak{C}$ la longueur de $s(t)$. On peut prendre $t_1 = 0$ et $t_n = T$) ou encore :

$$K(t) = \int_0^{\mathfrak{C}} [y(t+\tau) - s(\tau)]^2\, d\tau + \int_0^{t} y^2(\tau)\, d\tau + \int_{t+\mathfrak{C}}^{T} y^2(\tau)\, d\tau$$

Dans ce dernier cas, on a :

$$K(t) = \int_0^{\mathfrak{C}} s^2(\tau)\, d\tau + \int_0^{T} y^2(\tau)\, d\tau - 2\int_0^{\mathfrak{C}} \varepsilon(t+\tau)\, s(t - t_0 + \tau)\, d\tau - 2\int_0^{\mathfrak{C}} s(t - t_0 + \tau)\, s(\tau)\, d\tau$$

$K(t)$ sera minimal lorsque la corrélation mobile de $y(t)$ avec $s(t)$ sera maximale, c'est-à-dire en $t = t_0$.

Si plusieurs signaux $s_i(t - t_{0i})$ sont répartis au hasard sur O, T, on peut calculer les corrélations mobiles successives par les $s_i(t)$. On identifie le point de départ d'une impulsion par l'abscisse où la corrélation mobile de $s_i(t)$ avec $y(t)$ passe par un maximorum.

Cette méthode est une application simple de méthodes plus générales récentes (5) (15) (16) (19 *bis*) (22).

Si $s(t)$ est inconnu, on ne peut pratiquement faire qu'un examen à vue de $y(t)$.

b. Recherche de t_0 par minimisation des écarts quadratiques.

On peut rechercher la valeur t_0 telle que :

$$\mathcal{I}(t) = \int_0^{T} [y(\tau) - s(\tau - t)]^2\, d\tau \text{ soit minimal}$$

$$= \int_0^T [\varepsilon(\tau) + s(\tau - t_0) - s(\tau - t)]^2 \, d\tau$$

$$= 2\int_0^{\mathfrak{T}} s^2(\tau)\, d\tau + \int_0^T \varepsilon^2(\tau)\, d\tau - 2\int_0^{\mathfrak{T}-t_0+t} s(\tau)\, s(\tau + t_0 - t)\, d\tau + 2\int_0^T \varepsilon(\tau)\,[s(\tau - t_0) - s(\tau - t)]\, d\tau$$

Dans la mesure où les intercorrélations entre $\varepsilon(t)$ et $s(\tau - t)$ sont négligeables, on voit que $\mathcal{I}(t)$ est minimal pour $t = t_0$.

Pour minimiser $\mathcal{I}(t)$, il faut maximiser la corrélation mobile de $y(t)$ par $s(t)$, résultat déjà trouvé plus haut.

Dans le cas où t_0 a une densité de probabilité variable sur O, T, les résultats précédents ne sont plus valables, car il faut tenir compte des probabilités a priori, indiquées par $f(\tau)$, de présence de $s(t - t_0)$ en un point t_0.

La précision de la méthode de la corrélation mobile maximale est fonction de l'intercorrélation de $\varepsilon(t)$ et de $s(t)$. Si la probabilité que cette intercorrélation dépasse $\int_0^{\mathfrak{T}} s^2(t)\, dt$ reste faible, la méthode est satisfaisante. Ce problème sera repris plus loin à propos des spectres de $\varepsilon(t)$ et de $s(t)$.

c. Détection de $s[a(t + b)]$, $s(t)$ étant connu.

Si le milieu de propagation d'un signal modifie la période, la détection par la méthode de corrélation précédente de l'origine b du signal $s(t)$ connu peut encore s'appliquer, si a ne diffère pas trop de l'unité.

Posons $r(a, b) = \int_0^{\mathfrak{T}} s(t)\, s[a(t + b)]\, dt$:

si $a = 1$, $r(a, b)$ est l'autocorrélation de $s(t)$;

si $|b| > \frac{\mathfrak{T}}{a}$, si $a \to 0$, ou si $a \to \infty$, $r(a, b) = 0$

Localement $r(1, h)$ est maximal en $h = 0$ et $r(a, 0)$ est maximal en $a = 1$.

Soit $S_{a,b}(f)$ le spectre de $s[a(t + b)]$:

$$S_{a,b}(f) = \int_{-b}^{-b+\frac{\mathfrak{T}}{a}} s[a(t + b)]\, e^{-2i\pi ft}\, dt = \frac{1}{a}\int_0^{\mathfrak{T}} s(v)\, e^{-2i\pi \frac{f}{a}(v - ab)}\, dv$$

$$S_{a,b}(f) = \frac{1}{a} e^{2i\pi fb} S\left(\frac{f}{a}\right) \qquad \text{avec} \qquad S(f) = \text{spectre de } s(t)$$

L'autocorrélation de $s[a(t + b)]$ est :

$$\rho_a(h) = \int_0^{\frac{\mathfrak{T}-b}{a}} s(at)\, s[a(t + h)]\, dt = \frac{1}{a}\int_0^{\mathfrak{T}-b} s(u)\, s(ah + u)\, du$$

$$\rho_a(h) = \frac{1}{a}\rho_s(ah) \qquad \text{où} \qquad \rho_s(h) = \int_0^{\mathfrak{T}-h} s(t)\, s(t + h)\, dt$$

En étudiant par secteur l'autocorrélation d'une trace d'un sismogramme, on constate parfois un allongement de la période semblable à celui indiqué par la formule si $a < 1$.

2. Étude des procédés de mixage.

Lorsqu'on dispose, comme c'est le cas en sismique, de n enregistrements $y_i(t) = ks(t) + \varepsilon_i(t)$ où un même signal est reçu de façon systématique mais où ce signal est brouillé par des bruits qui diffèrent selon les enregistrements ($1 \leqslant i \leqslant n$), on a intérêt à calculer la moyenne des $y_i(t)$, car la variance des $\varepsilon_i(t)$ est divisée par n, alors que $ks(t)$ se retrouve systématiquement en t. La méthode suppose cependant une parfaite mise en phase de $s(t)$ sur tous les enregistrements. Sinon, la moyenne éliminera $ks(t)$ comme $\varepsilon(t)$. Un bruit $\varepsilon_i(t)$ sinusoïdal dont la phase varie aléatoirement selon les enregistrements est également éliminé par ce procédé de mixage, qui peut être effectué automatiquement par des machines à partir d'enregistrements magnétiques.

Comparaison des mixages additifs et multiplicatifs.

Supposons une trace composée d'un signal S_0 et d'un bruit ε de moyenne nulle et de loi $f(\varepsilon)$ avec une variance σ^2. Convenons d'appeler R le rapport signal sur bruit $R = \frac{S_0^2}{\sigma^2}$, rapport des énergies du signal S_0 et des énergies du bruit dans la trace. On peut facilement montrer l'apport d'un mixage lorsqu'on dispose de n traces i(1 à n) composées du même signal S_0 et de bruits ε_i de même loi $f(\varepsilon)$, indépendants entre eux.

Mixage additif :

$$y = S_0 + \frac{\sum_{i=1}^{n} \varepsilon_i}{n}$$

On a :

$$y^2 = S_0^2 + 2\frac{S_0}{n}\sum_i \varepsilon_i + \frac{1}{n^2}\sum_i \varepsilon_i^2$$

$$E(y^2) = S_0^2 + \frac{1}{n}\sigma^2$$

en espérance mathématique, ou lorsque $n \to \infty$.

Donc :

$$R_A = \frac{S_0^2}{\frac{\sigma^2}{n}} = nR$$

Le rapport signal/bruit est amélioré proportionnellement au nombre de traces mixées. Si nous opérons sur l'amplitude moyenne, on aura : $R_A = \sqrt{n}R$ avec le rapport de S_0 à σ.

Mixage multiplicatif :

$$y = \prod_{i=1}^{n} (S_0 + \varepsilon_i)$$

$$y^2 = \prod_{i=1}^{n} (S_0^2 + 2\varepsilon_i S_0 + \varepsilon_i^2) = S_0^{2n} + S_0^{2n-2}\sum_i \varepsilon_i^2 + S_0^{2n-4}\sum_{i \neq j} \varepsilon_i^2 \varepsilon_j^2 + \ldots + \varepsilon_1^2 \varepsilon_2^2 \ldots \varepsilon_n^2$$

Les termes de degré impair en ε_i ont une espérance mathématique nulle, de sorte que :

$$E[y^2] = S_0^{2n}\left[1 + \frac{n\sigma^2}{S_0^2} + \frac{n(n-1)}{2}\frac{\sigma^4}{S_0^4} + \ldots + C_n \frac{\sigma^{2j}}{S_0^j} + \ldots + \frac{\sigma^{2n}}{S_0^{2n}}\right]$$

$$= S_0^{2n}\left[1 + \frac{n}{R} + \ldots + C_n^j \frac{1}{R^j} + \ldots + \frac{1}{R^n}\right]$$

On aura donc : $$R_M = \frac{S_0^{2n}}{S_0^{2n}\left[1 + \frac{1}{R}\right]^n - S_0^{2n}} = \frac{1}{\left(1 + \frac{1}{R}\right)^n - 1}$$

Le rapport signal/bruit tend vers zéro. Dans le cas le plus favorable où R est très grand, on a :

$$R_M = \frac{1}{\frac{n}{R}} = \frac{R}{n}$$

Le mixage multiplicatif semble faire perdre de l'information, contrairement aux indications de K. Dyk (12) qui considère en fait des bruits de même fréquence, ce qui semble assez peu probable, et qui considère le rapport de l'amplitude moyenne du signal à l'amplitude moyenne du bruit. Ce rapport paraît peu pratique quand le bruit est d'amplitude moyenne nulle. D'autre part, Dyk considère le rapport du signal pur au bruit pur et fait l'hypothèse que les intercorrélations bruit et signal apportent une information à la fois sur le signal et sur le bruit. Si l'on considère que ces intercorrélations ne peuvent que brouiller un peu plus le signal, on voit que le rapport signal/bruit le plus représentatif serait le rapport :

$$\frac{S_0}{(S_0 + \varepsilon)^2} = \frac{\text{énergie du signal pur}}{\text{énergie du signal brouillé}}$$

rapport calculé après le mixage.

IV. NOMBRE DE RÉFLEXIONS, FILES D'ATTENTE, MULTIPLES

1. Nombre de réflexions en *t*.

On a écrit au paragraphe 4 du chapitre I que l'amplitude d'une trace sans le bruit était :

$$\int_{t-\mathfrak{T}}^{t} k\left(\frac{\tau}{2}\right) s\,(t-\tau)\,d\tau$$

On voit que le nombre de réflexions intervenant dans $y\,(t)$ à un instant t est :

$$\int_{t-\mathfrak{T}}^{t} k'\left(\frac{\tau}{2}\right) d\tau$$

où l'on peut considérer $k'\left(\frac{\tau}{2}\right)$, dérivée de $k\left(\frac{\tau}{2}\right)$ comme étant une suite de Dirac (arrivées de ré-

flexions $k\left(\frac{\tau}{2}\right)$ observables durant un temps $\mathfrak{T}$ à des instants τ). Si nous considérons que ces réflexions sont réparties à la Poisson avec une densité variable $\sigma\left(\frac{\tau}{2}\right)$, on sait que le nombre de réflexions présentes en t est $\int_{t-\mathfrak{T}}^{t} \sigma\left(\frac{\tau}{2}\right) d\tau$ en moyenne et que le nombre réel de réflexions suit une loi de Poisson autour de cette moyenne.

On trouve la même propriété dans les problèmes de files d'attente. Des clients arrivent selon un schéma de Poisson (indépendance entre les arrivées), mais avec une densité variable dans le temps : $\sigma(t)$. Ils restent dans le système de service durant un temps variable, régi par la loi F (t), probabilité qu'un service dure plus que t. Les clients peuvent aussi être des individus naissant en t et dont la loi de survie est F (t). On montre ci-dessous que le nombre d'individus présents à l'instant t est régi par une loi de Poisson dont la moyenne est :

$$m(t) = \int_{-\infty}^{t} \mathrm{F}(t-\tau)\,\sigma(\tau)\,d\tau$$

Dans le cas des réflexions, on a F $(t) = 1$ si $t < \mathfrak{T}$ et 0 ailleurs.

Démonstration : la présence à l'instant t d'un individu arrivé entre u et $u + du$ dans le service peut être notée par une variable aléatoire X égale à 1, et le fait qu'aucun individu n'est arrivé entre u et $u + du$, ou qu'il est déjà parti à l'instant t, peut être noté par une variable aléatoire X égale à 0.

La fonction caractéristique de X est :

$$\varphi_{u,t}(z) = 1 - \sigma(u)\,\mathrm{F}(t-u)\,du + e^{tz}\,\sigma(u)\,\mathrm{F}(t-u)\,du$$

La deuxième fonction caractéristique de X est :

$$\psi_{u,t}(z) = (e^{tz} - 1)\,\sigma(u)\,\mathrm{F}(t-u)\,du$$

La deuxième fonction caractéristique relative à la somme de tous les individus arrivés de $-\infty$ à t et encore présents dans le système en t est :

$$\psi_t(z) = (e^{tz} - 1) \int_{-\infty}^{t} \sigma(u)\,\mathrm{F}(t-u)\,du$$

La formule est exacte du fait de l'indépendance entre les arrivées des individus en u. Elle serait encore exacte si F $(t - u)$ était fonction de u et de $t - u$ (loi de temps de service variable en fonction de la date d'arrivée).

La fonction caractéristique de la variable aléatoire que constitue le nombre d'individus présents en t est :

$$\varphi_t(z) = e^{(e^{tz}-1)\int_{-\infty}^{t} \sigma(u)\,\mathrm{F}(u,t-u)\,du} = e^{(e^{tz}-1)\,m(t)}$$

où $m(t)$ est le nombre moyen d'individus présents en t. Cette forme de fonction caractéristique est la fonction caractéristique d'une variable aléatoire suivant une loi de Poisson de valeur moyenne :

$$m(t) = \int_{-\infty}^{t} \sigma(u)\,\mathrm{F}(u,t-u)\,du$$

La probabilité d'observer j individus présents en t est :

$$P(j,t) = e^{-m(t)} \frac{m^j(t)}{j!}$$

Remarquons que, dans le cas de clients demandant un service de loi F (t), nous supposons implicitement qu'il y a une infinité de postes de service en parallèle, prêts à servir chaque client dès son arrivée. La probabilité que M ou plus de M individus soient présents est :

$$\Theta(M, t) = 1 - \sum_{j=0}^{M-1} e^{-m(t)} \frac{m^j(t)}{j!} = 1 - Q(\chi^2, \nu)$$

avec :

$$m(t) = \frac{\chi^2}{2} \qquad \text{et} \qquad M = \frac{\nu}{2}$$

$$Q(\chi^2, \nu) = 2^{-\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)^{-1} \int_{\chi^2}^{\infty} e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1} dx$$

$Q(\chi^2, \nu)$ est en particulier tabulé dans *Biometrika* (6) [1]. Dans un article publié dans le n° 17 de la *Revue de la SOFRO*, nous donnons une application de la propriété précédente pour des calculs d'attente en se basant sur certaines approximations concernant des services à nombre fini de postes et des services à nombre infini de postes.

Des théorèmes concernant les probabilités d'attente dans des systèmes à nombre fini de serveurs, avec des lois de service quelconque, mais des taux d'arrivée à la Poisson constants, ont d'autre part été établis par FORTET et SEVASTYANOV (13) (26).

2. Réflexions multiples.

Les coefficients $k\left(\frac{\tau}{2}\right)$ envisagés jusqu'à présent correspondent à des coefficients provenant d'une réflexion simple située sur la couche rencontrée après un parcours de durée $\frac{\tau}{2}$ en partant de la surface à la vitesse du son dans le sol. Mais il faut également tenir compte des réflexions multiples parvenant à la surface un temps t après l'explosion. L'importance de ces réflexions pose un problème théoriquement mal résolu.

A un instant t, le sismographe enregistre les réflexions provenant des couches $\frac{t}{2}$ à $\frac{t-\mathcal{T}}{2}$ ($\mathcal{T}$: durée de l'impulsion). C'est le filtrage de ces coefficients par l'impulsion qui forme le signal du sismogramme. Mais il y a lieu de considérer d'autres coefficients de réflexion. En effet, pour une durée de parcours $2n$, et en supposant que nous avons un coefficient de réflexion à chaque pas : $k(1) \ldots k(n) \ldots$, on enregistre bien en surface au temps $2n$ le coefficient $k(n)$ — en négligeant les coefficients de transmission — mais on enregistre aussi toutes les réflexions multiples.

Sur la figure 13, on a figuré des chemins à 45° par rapport à l'horizontale pour mieux les séparer. En fait, les chemins suivent la verticale et l'origine et l'arrivée sont confondues, mais cette représentation permet de mieux différencier les chemins des réflexions multiples.

Sur cette figure, nous avons tracé un chemin possible utilisant cinq réflexions successives et reve-

[1] Vol. I, p. 122-134.

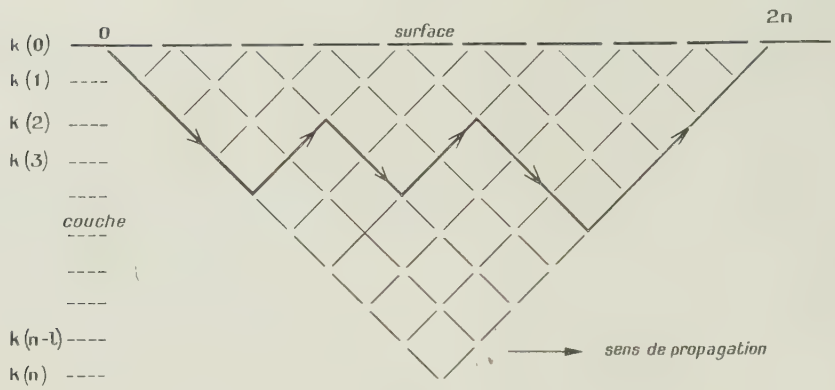


FIG. 13.

nant en surface au temps $2n$. Il y a des réflexions d'ordre 1, 3, 5, 7 ..., $2n - 1$ parvenant en surface au temps $2n$. Il y a une réflexion d'ordre 1 et une réflexion d'ordre $2n - 1$. On montre facilement qu'il y a $\frac{n(n-1)}{2}$ réflexions d'ordre 3 et autant d'ordre $2n - 3$. On peut montrer que le nombre de réflexions d'ordre 5 est :

$$N_n^5 = \frac{n(n-1)^2(n-2)}{12}$$

On peut induire que la formule générale est :

$$N_n^{2h+1} = \frac{n(n-1)^2(n-2)^2 \dots (n-h+1)^2(n-h)}{1 \cdot 2^2 \cdot 3^2 \dots h^2(h+1)}$$

$$N_n^{2h+1} = \frac{C_n^h C_{n-1}^h}{h+1}$$

3. Nombre de réflexions multiples d'ordre 2 h + 1.

Nous figurons ici les chemins suivis par les ondes à 45° par rapport à la surface pour distinguer les réflexions multiples. On peut déterminer le chemin suivi par une réflexion multiple

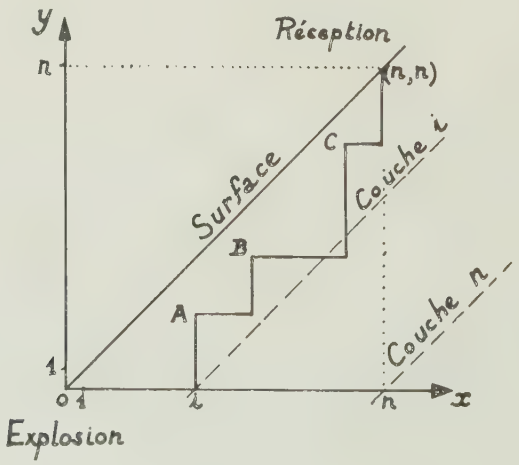


FIG. 13 *bis*.

d'ordre $2h+1$ revenant en surface à l'instant $2n$ après l'explosion de l'instant 0, en fixant les positions (x, y) qui sont des valeurs entières, des h sommets A, B, C... du chemin.

Pour déterminer le nombre de réflexions multiples d'ordre $2h+1$, soit $N^{2h+1}_{X,Y}$ remontant en un point (X, Y) dans le sol ($X \geqslant Y$), on utilise la relation suivante, qui nous a été suggérée par NEVEU :

$$N^{2h+1}_{X,Y} = \sum_{\substack{h \leqslant y \leqslant Y-1 \\ y \leqslant x \leqslant X-1}} N^{2h-1}_{x,y} \qquad (x \leqslant y)$$

Il est évident que $N^{1}_{X,Y} = 1$:

$$N^{3}_{X,Y} = \sum_{\substack{1 \leqslant y \leqslant Y-1 \\ y \leqslant x \leqslant X-1}} 1 = \frac{(Y-1)(2X-Y)}{2}$$

$$N^{5}_{X,Y} = \sum_{\substack{2 \leqslant y \leqslant Y-1 \\ y \leqslant x \leqslant X-1}} \frac{(y-1)(2x-y)}{2} = \frac{(X-1)(Y-1)(Y-2)(3X-2Y)}{2!\,3!}$$

$$N^{7}_{X,Y} = \frac{(X-1)(X-2)(Y-1)(Y-2)(Y-3)(4X-3Y)}{3!\,4!}$$

On démontre ces relations en faisant les sommes des x ou y entiers, de leurs carrés et de leurs cubes.

On peut déduire des résultats précédents que :

$$\boxed{N^{2h+1}_{X,Y} = C^{h}_{X} C^{h}_{Y-1} - C^{h-1}_{X-1} C^{h+1}_{Y}}$$

La formule étant vraie pour $h = 0, 1, 2$ et 3 et étant admise pour h, on peut la démontrer par récurrence pour $h+1$:

$$N^{2h+3}_{X,Y} = \sum_{\substack{h+1 \leqslant y \leqslant Y-1 \\ y \leqslant x \leqslant X-1}} [C^{h}_{x} C^{h}_{y-1} - C^{h-1}_{x-1} C^{h+1}_{y}]$$

$$= \sum_{h+1 \leqslant y \leqslant Y-1} \{ C^{h}_{y-1} [C^{h}_{X-1} + \ldots + C^{h}_{y}] - C^{h+1}_{y} [C^{h-1}_{X-2} + \ldots + C^{h-1}_{y-1}] \}$$

$$= \sum_{h+1 \leqslant y \leqslant Y-1} \{ C^{h}_{y-1} [C^{h+1}_{X} - C^{h+1}_{y}] - C^{h+1}_{y} [C^{h}_{X-1} - C^{h}_{y-1}] \}$$

car :

$$C^{h}_{X-1} + \ldots + C^{h}_{y} + \ldots + C^{h}_{h} = C^{h+1}_{X}$$

$$N^{2h+3}_{X,Y} = \sum_{h+1 \leqslant y \leqslant Y-1} [C^{h}_{y-1} C^{h+1}_{X} - C^{h+1}_{y} C^{h}_{X-1}]$$

$$= C^{h+1}_{X} [C^{h}_{Y-2} + \ldots + C^{h}_{h}] - C^{h}_{X-1} [C^{h+1}_{Y-1} + \ldots + C^{h+1}_{h+1}] = C^{h+1}_{X} C^{h+1}_{Y-1} - C^{h}_{X-1} C^{h+2}_{Y}$$

On vérifie que $N^{2h+3}_{X,Y}$ a bien la forme générale et la formule est ainsi démontrée.

Lorsque $X = Y = n$ (réception en surface au temps $2n$) nous avons :

$$N^{2h+1}_{n} = C^{h}_{n} C^{h}_{n-1} - C^{h-1}_{n-1} C^{h+1}_{n} = \frac{C^{h}_{n} C_{n-1}}{h+1}$$

Donc on trouve :

$$N_n^{2h+1} = \frac{C_n^h C_{n-1}^h}{h+1} = \frac{C_n^h C_n^{h+1}}{n} \quad (1)$$

Le tableau suivant indique les valeurs de N_n^{2h+1} pour quelques valeurs de h et de n.

n	N_n^3	N_n^5	N_n^7	N_n^9	N_n^{11}	N_n^{13}	N_n^{15}
2	1						
3	3	1					
4	6	6	1				
5	10	20	10	1			
6	15	50	50	15	1		
7	21	105	175	105	21	1	
8	28	196				28	1
	$\frac{n(n-1)}{2}$	$\frac{n(n-1)^2(n-2)}{2^2 \cdot 3}$					

Sur la formule $N_n^{2h+1} = \frac{C_n^h C_{n-1}^h}{h+1}$, on voit que N_n^{2h+1} croît avec h jusqu'en $h \leqslant \frac{n-3}{2}$. La croissance est rapide ($N_n^1 = 1$, $N_n^3 \simeq \frac{n^2}{2}$, $N_n^5 \simeq \frac{n^4}{12}$ ...)

Si l'on admet que tous les chemins sont indépendants, de même que les coefficients de réflexion qu'ils utilisent, ce qui revient à minimiser leur contribution globale, on sait que la loi de la somme des coefficients des réflexions d'ordre $2k+1$ sera une loi normale de moyenne nulle et de variance. $\sigma^2(2k+1) = N^{2k+1}\sigma^{2(2k+1)}$. La contribution de toutes les réflexions multiples ou non provenant au temps $2n$ suivra une loi normale de moyenne nulle et de variance :

$$\sigma_i^2 = \sigma^2 + \frac{n(n-1)}{2}\sigma^6 + \ldots + N_n^{2k+1}\sigma^{2(2k+1)} + \ldots + \sigma^{2(2n-1)}$$

Cette variance totale σ_i^2 est très supérieure à σ^2 de sorte que les réflexions multiples étoufferaient les réflexions simples.

En réalité, les choses sont plus complexes car les réflexions multiples ne sont pas indépendantes. On peut retrouver les mêmes coefficients de réflexion sur des chemins différents. Les coefficients correspondant à ces chemins ne sont pas indépendants. Si nous le supposions, on voit par exemple que les réflexions d'ordre 3 parvenant en surface au temps $2n$ fourniraient une contribution :

$$\sum_{i=1}^{n-1}\sum_{j=0}^{i-1} k(i)\,k(j)\,k(n-1+j) \quad (2)$$

En négligeant les $\frac{n}{2}$ cas où $k(i) = k(n-i+j)$, il reste $\frac{n^2}{2}$ chemins qu'on peut grouper deux à

(1) On suppose ici, pour simplifier, les coefficients de transmission égaux à 1 et non à $1 - k(i)$.

(2) *Remarque* : Dans un article de *Izv. Akad. Nauk S. S. S. R. ser. geof.*, Moskva, 1959, n° 8 (1089-1102), A. M. Epinat'eva donne sans démonstration une formule de calcul de N_n^{2k+1} par récurrence différente de la nôtre. Le tableau des valeurs de N_n^{2k+1} diffère du nôtre pour la plupart des valeurs et nous avons pu montrer que les formules et le tableau proposés par Epinat'eva étaient faux.

deux car les rôles des $k\,(i)$ et $k\,(n-i+j)$ y sont symétriques. On a alors $\frac{n^2}{4}$ variables indépendantes de forme $2\,k\,k'k''$ et leur somme suit une loi de Gauss de moyenne nulle comme E (k) et de variance $\frac{n^2}{4}\,4\,\sigma^6 = n^2\sigma^6$. Si n est assez grand, on voit que la variance de la contribution des multiples d'ordre 3 peut dépasser la variance σ^2 des réflexions simples.

Ceci semblerait condamner les méthodes sismiques, puisque l'influence de toutes les couches réfléchissantes se mélangerait. En fait, il n'en est pas ainsi. Les multiples ne sont gênants que dans certains terrains à forte couche réfléchissante superficielle et vers la fin des films où leur influence s'accroît théoriquement et où l'absorption interdit toute interprétation.

En utilisant une méthode matricielle et en opérant sur les spectres, GRAU construit directement le sismogramme synthétique à multiples correspondant à une série de $k\,(i)$ donnés, sans omettre les facteurs de transmission (2) et (8). Les sismogrammes synthétiques avec et sans multiples ne diffèrent pratiquement pas au début, mais l'importance des multiples devient de plus en plus grande lorsque t croît. Au-delà d'une ou deux secondes, les multiples peuvent masquer les réflexions simples.

L'équation du signal utilisée dans nos calculs suppose le sismogramme constitué uniquement de réflexions simples sans coefficient de transmission. Pour obtenir l'équation analogue du sismogramme à multiples, il faut remplacer $k\left(\frac{\tau}{2}\right)$ par le coefficient $k_m\left(\frac{\tau}{2}\right)$ qui contient l'influence de tous les coefficients de réflexion (simple et multiples corrigés par les coefficients de transmission) revenant en surface à l'instant τ.

V. ANALYSE HARMONIQUE, AUTOCORRÉLATION ET INTERCORRÉLATION

1. Intérêt de ces calculs.

On a vu plus haut que les coefficients de réflexion $k\,(t)$ peuvent être supposés répartis à la Poisson au cours du temps et que leurs amplitudes suivent approximativement une loi de Gauss. En fait, ces indications sont valables sur un secteur d'une demi à une seconde, mais la densité des $k\,(t)$ et les paramètres de leurs amplitudes (moyenne et écart type) peuvent varier d'un secteur à l'autre. De plus, l'énergie revenant en surface à la suite d'un tir décroît de 1 à quelques millièmes entre le début et la fin de l'enregistrement du fait de l'absorption. En dépit de contrôles de gain destinés à maintenir constante l'énergie enregistrée par les sismographes, il peut y avoir des écarts d'énergie du simple au double d'un secteur à l'autre. Cependant, sur un secteur d'une seconde, on peut supposer le signal $y\,(t)$ approximativement stationnaire. La répartition des $k\,(t)$ à la Poisson permet alors d'éliminer pratiquement leur influence dans des calculs d'analyse harmonique et d'autocorrélation et de retrouver les caractéristiques de l'impulsion $s\,(t)$ inconnue a priori.

2. Analyse harmonique.

On suppose le sismogramme connu de O à T et on le soumet à l'analyse harmonique sur un sous-secteur T_i à T_{i+1}.

$$Y\,(f) = \int_{T_i}^{T_{i+1}} y\,(t)\,e^{-2i\pi ft}\,dt = \int_{T_i}^{T_{i+1}} e^{-2i\pi ft} \int_{t-\bar{\omega}}^{t} k\left(\frac{\tau}{2}\right) s\,(t-\tau)\,d\tau dt + G\,(f) + E\,(f)$$

où G (*f*) et E (*f*) sont les spectres des composantes $g(t)$ et du bruit $\varepsilon(t)$ sur l'intervalle T_i, T_{i+1}. Si l'intégrale double portait sur des intervalles de variation infinis, le spectre du produit de convolution de k par s serait égal au produit des spectres de k et de s. Sur des secteurs finis, ce résultat n'est qu'approximatif.

On a :

$$J = \int_{i}^{T_{i+1}} \int_{t-\mathfrak{C}}^{t} s(t-\tau)\, k\left(\frac{\tau}{2}\right) e^{-2i\pi ft}\, d\tau\, dt$$

Le domaine de variation de t et τ est la zone hachurée de la figure 14. Nous pouvons la décomposer en trois domaines, comme indiqué. Il correspond trois intégrales J_1, J_2 et J_3 à ces domaines. Dans les intégrales J_1 et J_3, les variables interviennent dans les bornes des domaines d'intégration et la séparation des spectres de s et k n'est pas possible. En revanche, on peut intégrer dans le domaine 2.

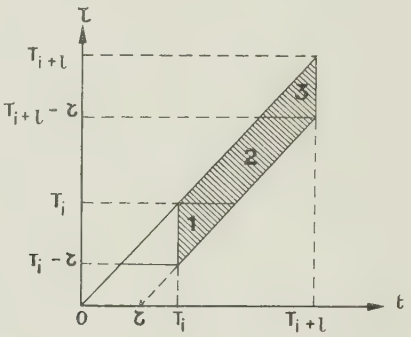


FIG. 14.

Les intervalles de variation de t et τ dans le domaine 2 sont :

$$\begin{array}{lllll} T_i & < & \tau & < & T_{i+1} - \mathfrak{C} \\ \tau & < & t & < & \tau + \mathfrak{C} \quad \text{pour } \tau \text{ fixé} \end{array}$$

On peut donc écrire :

$$J_2 = \int_{\tau}^{\tau+\mathfrak{C}} \int_{T_i}^{T_{i+1}-\mathfrak{C}} k\left(\frac{\tau}{2}\right) s(t-\tau)\, e^{-2i\pi ft}\, d\tau\, dt$$

Effectuons un changement de variable $u = t - \tau$ dans l'intégration en t :

$$J_2 = \int_{T_i}^{T_{i+1}-\mathfrak{C}} \int_{0}^{\mathfrak{C}} k\left(\frac{\tau}{2}\right) e^{-2i\pi ft}\, s(u)\, e^{-2i\pi fu}\, du\, d\tau = K(f) \cdot S(f)$$

car les deux intégrales se séparent. K (*f*) est le spectre de k entre T_i et $T_{i+1} - \mathfrak{C}$ et S (*f*) est le spectre de $s(t)$. Dans la mesure où J_1 et J_3 ne sont pas négligeables devant J_2, l'approximation de J par J_2 n'est pas convenable.

L'approximation est d'autant plus satisfaisante que le secteur T_i, $T_{i+1} - \mathfrak{C}$ est long, mais en ce cas, $k(t)$ ne restant pas stationnaire sur un très long secteur et $s(t)$ se déformant également, K (*f*) et S (*f*) se modifient.

En général, si G (f), E (f), K (f) et S (f) ne sont pas connus, la connaissance de Y (f) ne permet pas de préciser $s\,(t)$ et $k\,(t)$. Cependant, si l'on suppose que le signal $y\,(t)$ ne comporte pas de bruit et si $k\,(t)$ est une suite de coefficients de réflexion discrets à la Poisson, le spectre K (f) calculé sur un secteur assez long est constant et le spectre de $y\,(t)$ se réduit à :

$$Y\,(f) = K \cdot S\,(f)$$

Soient $\mathcal{R}\,(f)$ et $\mathcal{I}\,(f)$ les parties réelles et imaginaires de S (f). Comme $s\,(t)$ est nul de $-\infty$ à o, on a :

$$s\,(t) = \int_{-\infty}^{\infty} [\mathcal{R}\,(f)\,\cos\,2\pi ft - \mathcal{I}\,(f)\,\sin\,2\pi ft]\,df = 0 \text{ pour tout } t < 0$$

Lorsque t change de signe, on a donc :

$$\int_{-\infty}^{\infty} \mathcal{R}\,(f)\,\cos\,2\pi ft\,df = -\int_{-\infty}^{\infty} \mathcal{I}\,(f)\,\sin\,2\pi ft\,df$$

On peut donc écrire :

$$s\,(t) = 4\int_{0}^{\infty} \mathcal{R}\,(f)\,\cos\,2\pi ft\,df$$

Remarquons que l'utilisation de l'analyse harmonique, pour reconstituer le signal $y\,(t)$ à partir de Y (f), nécessite la connaissance de Y (f) avec une densité d'échantillonnage très serrée :

$$y\,(t) = \int_{-\infty}^{\infty} Y\,(f)\,e^{2i\pi ft}\,df = \lim_{\Delta f \to 0} \Delta f \sum_{n=-\infty}^{\infty} Y\,(n\,\Delta f)\,e^{2i\pi n\,\Delta f t}$$

Si Y (f) n'est pas connu avec un échantillonnage très fin, lorsqu'on calcule $y\,(t)$ pour $t > A$, on risque de commettre de grandes erreurs, car les valeurs $nt\,\Delta f$ fournissent un échantillonnage insuffisant de $e^{2i\pi nt\Delta f}$. Il faut que $\Delta f \ll \frac{1}{A}$ pour avoir de bons résultats. Inversement, la connaissance précise de Y (f) pour les grandes valeurs de f suppose un échantillonnage serré de $y\,(t)$, pour la même raison.

3. Autocorrélation.

D'une manière générale, l'autocorrélation est beaucoup utilisée par les statisticiens en sismique. En fait, elle présente à peu près les mêmes inconvénients que ceux de l'analyse harmonique, ce qui s'explique bien puisque le spectre de l'autocorrélation est égal au carré de Y (f).

Les fonctions $y\,(t)$ considérées ici sont évidemment du second ordre : E $[|y\,(t)|^2]$ existe. De ce fait, d'après l'inégalité de Schwartz, leurs covariances E $[y\,(t)\,\overline{y\,(t')}]$ existent.

Les fonctions $y\,(t)$ sont des fonctions aléatoires stationnaires d'ordre 2 : en effet, elles sont :

1) stationnaires de premier ordre ;
2) leur autocorrélation est stationnaire ;
3) elles sont continues en moyenne quadratique.

1) On a : $$E[y\,(t)] = \lim_{T \to \infty} \frac{1}{T} \int_{t}^{t+T} y\,(t)\,dt = 0 \text{ quel que soit } t$$

2) $$E[y^2\,(t)] = \lim_{T \to \infty} \frac{1}{T} \int_{t}^{t+T} y^2\,(t)\,dt = C^{te} \text{ quel que soit } t$$

La forme de l'autocorrélation est :

$$\rho(h) = \mathrm{E}[y(t) \cdot \overline{y(t+h)}] = \lim_{\mathrm{T}\to\infty} \frac{1}{\mathrm{T}} \int_{t}^{t+\mathrm{T}} y(t)\, \overline{y(t+h)}\, dt$$

L'emploi des quantités complexes conjuguées $\overline{y(t)}$ est inutile dans le cas des sismogrammes qui sont des signaux réels.

Propriétés de l'autocorrélation avec 1), 2) et 3) (voir (7) et (19)) :

$\rho(h)$ est indépendant de t ;
D'après l'inégalité de SCHWARTZ, $\rho(h) \leqslant \rho(0)$;
$\rho(h)$ est une fonction paire. Son spectre $= \mathrm{Y}^2(f)$;

$$\lim_{\mathrm{T}\to\infty} \frac{1}{\mathrm{T}} \int_{0}^{\mathrm{T}} \rho(h)\, dh \to 0$$

Implicitement, nous supposons que $y(t)$ est connu de t à $t + \mathrm{T} + h$: sinon, la formule de l'autocorrélation serait :

$$\rho(h) = \frac{1}{\mathrm{T} - h} \int_{t}^{t+\mathrm{T}-h} y(t)\, y(t+h)\, dt$$

Cette dernière quantité nous semble moins homogène que la première en ce qui concerne la variance de $\rho(h)$.
En fait, nous utilisons ici des autocorrélations normées :

$$\rho_n(h) = \frac{\int_{t}^{t+\mathrm{T}} y(t)\, y(t+h)\, dt}{\sqrt{\int_{t}^{t+\mathrm{T}} y^2(t)\, dt \int_{t+h}^{t+\mathrm{T}+h} y^2(t)\, dt}}$$

La norme a pour effet de corriger en partie le manque de stationnarité de l'énergie, car, en fait, l'énergie varie un peu au cours de l'enregistrement et le signal $s(t)$ se déforme.

Calcul de $\rho(h)$ avec $y(t) = \int_{t-\mathfrak{T}}^{t} s(t-\tau)\, k\left(\frac{\tau}{2}\right) d\tau + \varepsilon(t)$:

$$\underset{(\mathrm{T}=\mathrm{T}_{i+1}-\mathrm{T}_i)}{\mathrm{E}[\rho(h)]} = \mathrm{E}\left\{ \frac{1}{\mathrm{T}} \int_{\mathrm{T}_i}^{\mathrm{T}_{i+1}} \int_{t-\mathfrak{T}}^{t} s(t-\tau)\, k\left(\frac{\tau}{2}\right) d\tau \int_{t+h-\mathfrak{T}}^{t+h} s(t-\tau'+h)\, k\left(\frac{\tau'}{2}\right) d\tau'\, dt \right.$$

$$+ \frac{1}{\mathrm{T}} \int_{\mathrm{T}_i}^{\mathrm{T}_{i+1}} \varepsilon(t)\, \varepsilon(t+h)\, dt + \frac{1}{\mathrm{T}} \int_{\mathrm{T}_i}^{\mathrm{T}_{i+1}} \varepsilon(t) \int_{t+h-\mathfrak{T}}^{t+h} s(t-\tau')\, k\left(\frac{\tau'}{2}\right) d\tau'\, dt$$

$$\left. + \int_{\mathrm{T}_i}^{\mathrm{T}_{i+1}} \varepsilon(t+h) \int_{t-\mathfrak{T}}^{t} s(t-\tau)\, k\left(\frac{\tau}{2}\right) d\tau\, dt \right\}$$

Dans la mesure où le bruit $\varepsilon(t)$ est sans autocorrélation et de moyenne nulle, les trois dernières intégrales donnent un résidu $\eta(h)$ de moyenne nulle.

Faisons un changement : $\begin{cases} t-\tau=u \\ t-\tau'+h=v \end{cases}$ dans le premier terme.

$$\mathrm{E}[\rho(h)] = \mathrm{E}\left\{ \eta(h) + \int_{T_i}^{T_{i+1}} \int_0^{\mathfrak{T}} \int_0^{\mathfrak{T}} k\left(\frac{t-u}{2}\right) k\left(\frac{t-v+h}{2}\right) s(u)\, s(v)\, du\, dv\, dt \right\}$$

Le fait de considérer $y(t)$ entre T_i et T_{i+1} suppose en réalité que nous utilisons $k(t)$ de $T_i - \mathfrak{T}$ à T_{i+1}. Dans l'intégrale précédente, si nous intégrons d'abord en t, nous ferons apparaître :

$$\rho_k\left(\frac{h-v+u}{2}\right) = \int_{T_i}^{T_{i+1}} k\left(\frac{t-u}{2}\right) k\left(\frac{t+h-v}{2}\right) dt$$

D'après les propriétés de stationnarité de y et de k, ce terme $\rho_k\left(\frac{h-v+u}{2}\right)$ est indépendant des bornes d'intégration :

$$\int_{T_i-u}^{T_{i+1}-u} k\left(\frac{t}{2}\right) k\left(\frac{t+h-v+u}{2}\right) dt = \int_{T_i-x}^{T_{i+1}-x} k\left(\frac{t}{2}\right) k\left(\frac{t+h-v+u}{2}\right) dt$$

On a donc :

$$\rho(h) = \eta(h) + \int_0^{\mathfrak{T}} \int_0^{\mathfrak{T}} \rho_k\left(\frac{h-v+u}{2}\right) s(u)\, s(v)\, du\, dv$$

Un nouveau changement en $x = v - u$ donne :

$$\rho(h) = \eta(h) + \int_0^{\mathfrak{T}} \int_{-u}^{\mathfrak{T}-u} s(u)\, s(u+x)\, \rho_k\left(\frac{h-x}{2}\right) du\, dx = \eta(h) + \int_{-\mathfrak{T}}^{\mathfrak{T}} \rho_k\left(\frac{h-x}{2}\right) \rho_s(x)\, dx$$

En effet, le domaine de variation de x et de u est le domaine hachuré (fig. 15), puisque $s(t) = 0$ lorsque t est extérieur à $(0, \mathfrak{T})$. On peut donc fixer x et faire varier u :

$$\begin{cases} \text{si } x > 0 & 0 < u < \mathfrak{T} - x \\ \text{si } x < 0 & -x < u < \mathfrak{T} \end{cases}$$

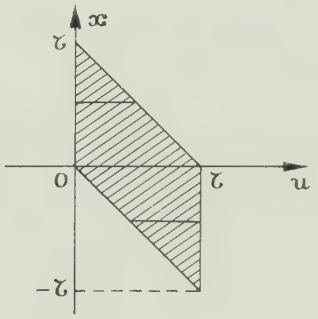


FIG. 15.

$$\int_0^{\mathfrak{T}-x} s(u)\, s(u+x)\, du = \rho_s(x) \left(= \int_{-x}^{\mathfrak{T}} s(u)\, s(u+x)\, du \qquad \text{avec} \qquad x < 0 \right)$$

Il faut alors faire varier x de $-\mathfrak{T}$ à $\mathfrak{T}$:

$$\rho(h) = \int_{-\mathfrak{T}}^{\mathfrak{T}} \rho_k \left(\frac{h-x}{2} \right) \rho_s(x)\, dx + \eta(h)$$

$\eta(h) \to 0$ quand T croît.

Lorsque les coefficients $k(t)$ sont répartis à la Poisson, $\rho_k \left(\frac{h-x}{2} \right)$ est une fonction de Dirac, nulle pour $h \neq x$.

De ce fait, on obtient :

$$\rho(h) = C\rho_s(h) + \eta(h)$$

En fait, si nous intégrons sur un ensemble T fini, $\rho_k \left(\frac{h-x}{2} \right)$ n'est pas nul pour $h \neq x$. On peut observer dans le cas particulier de la figure 8 (1) où T est de l'ordre d'une seconde, que $\rho_k(h)$ est souvent de l'ordre de 10 % de $\rho_k(0)$. D'autre part, lorsque h croît, l'intervalle d'intégration des $k(t)$ se déplaçant, on risque de modifier la valeur théoriquement nulle de $\rho_k \left(\frac{h-x}{2} \right)$ par l'introduction de coefficients k importants entre T_{i+1} et $T_{i+1} + h$. Pour amoindrir cet inconvénient, nous normons l'autocorrélation, comme indiqué p. 42.

4. Variance de l'autocorrélation.

Le calcul de l'autocorrélation n'est utile que dans la mesure où la variance de cette dernière n'est pas très importante. Nous avons vu au chapitre II, paragraphe 5, que les amplitudes $y(t)$ des sismogrammes suivaient une loi normale de moyenne nulle et d'écart type σ. Dans l'hypothèse où les amplitudes sont indépendantes, on peut dire que $\int_0^T y^2(t)\, dt$ a pour espérance mathématique $KT\sigma^2$. Comme on a $E[(y^2 - \sigma^2)^2] = 2\sigma^4$, on voit que la variance de l'autocorrélation pour $h = 0$ est $2\sigma^4 TK$. Les variations relatives de $\rho(0)$ selon le secteur de longueur T où on le calcule ont 95 % de chances d'être comprises entre :

$$\pm 2 \frac{\sqrt{2T\sigma^4 K}}{KT\sigma^2} = \pm 2 \sqrt{\frac{2}{TK}}$$

Tant que $h < \mathfrak{T}$, longueur de l'impulsion $s(t)$, les amplitudes $y(t)$ ne sont pas indépendantes, de sorte que la formule précédente est seulement approximative. Lorsque $h < -\mathfrak{T}$ et $h > \mathfrak{T}$, on a en théorie $E[y(t)\, y(t+h)] = 0$. Mais la variance de $y(t)\, y(t+h)$ est égale à $KT\sigma^4$. Les variations relatives de $\frac{\rho(h)}{\rho(0)}$ ont 95 % de chances d'évoluer entre $\pm 2 \sqrt{\frac{1}{TK}}$.

(1) Voir p. 18.

Le facteur K, déterminé par l'expérience tient compte de l'unité adoptée pour la mesure de T et de la dépendance existant entre les $y(t)$ pour $h < \mathcal{C}$.

En fait, la décroissance de l'écart type de $\rho(h)$ de la forme $2\sqrt{\frac{1}{TK}}$ semble un peu rapide. Un certain nombre d'expériences d'autocorrélations sur diverses traces et sur des secteurs de 500 ou 800 ms ont été faites dans des cas où les impulsions variaient de 30 à 80 ms selon les tirs. En ajoutant les autocorrélations de trois ou cinq secteurs adjacents d'une même trace (l'autocorrélation jouissant de la propriété d'additivité en cas de stationnarité), on obtient des autocorrélations sur des secteurs de 1 500, 2 400 et 2 500 ms.

Nous avons relevé les valeurs de ces autocorrélations toutes les 20 ms pour $h \geqslant 80$ ms, et nous avons étudié leur répartition.

L'ajustement des valeurs observées de $\rho(h)$ normé ($\rho(0) = 1$) à une loi normale de moyenne nulle est satisfaisant. La figure 16 montre les valeurs obtenues pour l'écart type σ_r correspondant aux divers ajustements. On observe bien une décroissance de σ_r quand le secteur T de calcul de $\rho(h)$ s'accroît, mais cette décroissance est moins rapide que la décroissance théorique en $\sqrt{\frac{1}{KT}}$.

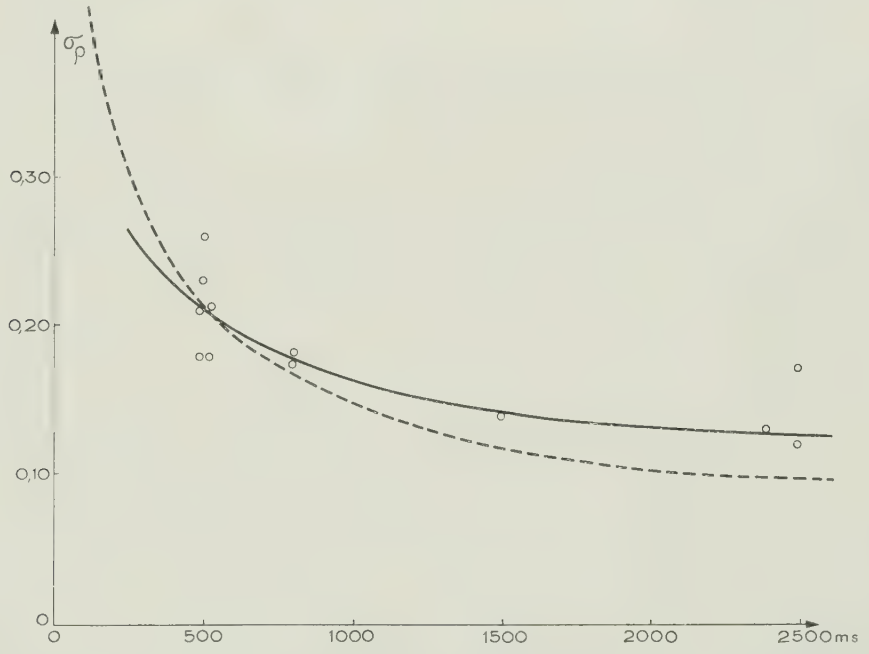


FIG. 16. — Décroissance de l'écart type de l'autocorrélation avec l'accroissement de T.

○ observations —— courbe ajustée ---- décroissance en $\frac{k}{\sqrt{T}}$.

5. Intercorrélation.

Pour rechercher la meilleure mise en phase de deux traces d'un même film, on peut calculer leur intercorrélation :

$$r_{i,j}(h) = \frac{1}{T}\int_0^T v_i(t)\, y_j(t+h)\, dt$$

Dans la mesure où un film n'est composé que de réflexions et où le couplage des sismographes des diverses traces avec le sol est satisfaisant, on devrait obtenir une intercorrélation normée :

$$ {}_n\gamma_{i,j}(h) = \frac{\int_0^T y_i(t)\, y_j(t+h)\, dt}{\sqrt{\int_0^T y_i^2(t)\, dt \int_0^T y_j^2(t+h)\, dt}} $$

très voisine de 1 pour le déphasage h_0 mettant les réflexions en concordance de phase sur les traces i et j. Sur les films que nous avons analysés, pour lesquels il n'était pas possible d'observer des réflexions à vue, l'intercorrélation ne permet pas de déterminer un déphasage maximisant ${}_n\gamma_{i,j}(h)$. Comme la variance de l'intercorrélation est au moins égale à celle de l'autocorrélation, les maximums de ${}_n\gamma_{i,j}(h)$ sont aussi bien imputables au hasard qu'à une mise en phase.

En théorie, on peut montrer que :

$$ {}_n\gamma_{i,j}(h) = \int_{-\infty}^{\infty} \rho_k \left(\frac{h - x + \varphi_i - \varphi_j}{2} \right) \rho_s(x)\, dx + \varepsilon(h) $$

$\varepsilon(h)$ étant la somme de l'intercorrélation des bruits entre eux et de leur intercorrélation avec les réflexions. On voit que ${}_n\gamma_{i,j}(h)$ passe par un maximum quand $h = \varphi_i - \varphi_j$ (déphasage entre les traces i et j).

6. Analyse d'une composante périodique.

Nous étudions ici la présence d'une composante périodique $g(t)$ dans la trace d'un sismogramme, en particulier une sinusoïde ou une sinusoïde à période variable (la propagation dans le sol se fait souvent avec une variation de la période).

a. Analyse harmonique

1) $g(t) = k\, e^{2i\pi\nu(t+\varphi)}$.

$$ Y(f) = \int_0^T y(t)\, e^{-2i\pi ft}\, dt = K(f) \cdot S(f) + E(f) + \int_0^T k e^{2i\pi\nu(t+\varphi) - 2i\pi ft}\, dt $$

$$ = K(f) \cdot S(f) + E(f) + \frac{k \sin \pi(\nu - f) T}{\pi(\nu - f)}\, e^{2i\pi\nu\varphi + i\pi(\nu - f)T} $$

Lorsque $f \to \nu$, $\dfrac{\sin \pi(\nu - f) T}{\pi(\nu - f)} \to T$ et l'on a :

$$ \boxed{Y(\nu) = S(\nu)\, K\left(\frac{\nu}{2}\right) + E(\nu) + T k e^{2i\pi\nu\varphi}} $$

Lorsque $\nu \neq f$, le spectre du sinus ne tend vers zéro que si T est très grand. L'analyse harmonique ne révèle l'existence de la sinusoïde que si T et k sont assez grands.

2) $g(t) = k\, e^{2i\pi\nu(l + ct)t}$.

Si $c > 0$, on a une compression de la période quand t croît.

Si $c < 0$, on a une dilatation de la période quand t croît.

L'analyse harmonique de $g(t)$ donne :

$$G(f) = k \int_0^T e^{2i\pi(\nu - f)t + 2i\pi\nu ct^2}\, dt$$

Si $f \neq \nu$ et $T \to \infty$, $G(f)$ est voisin de 0.
Si $f = \nu$, on obtient :

$$G(\nu) = k \int_0^T e^{2i\pi\nu ct^2}\, dt$$

$$= k \int_0^T [\cos 2\pi\nu ct^2 + i \sin 2\pi\nu ct^2]\, dt = \frac{k}{2\sqrt{\nu c}} \int_0^{2T\sqrt{\nu c}} \left(\cos \frac{\pi}{2} u^2 - i \sin \frac{\pi}{2} u^2\right) du$$

en posant $u = 2t\sqrt{\nu c}$.

$X(z) = \int_0^z \cos \frac{\pi}{2} u^2\, du$ et $Y(z) = \int_0^z \sin \frac{\pi}{2} u^2\, du$ sont les intégrales de Fresnel, de sorte que :

$$G(\nu) = \frac{k}{2\sqrt{\nu c}} \left[X\left(2T\sqrt{\nu c}\right) - iY\left(2T\sqrt{\nu c}\right)\right]$$

Lorsque z dépasse 2, $X(z)$ et $Y(z)$ restent toujours inférieurs à 0,77 et $G(\nu)$ restera inférieur à $\frac{kT}{3}$ de sorte que par l'analyse harmonique la contribution de $g(t)$ est réduite au tiers de ce qu'elle serait si c égalait zéro. Il faut donc que $2T\sqrt{\nu c} < 2$, ou $|c| < \frac{1}{\nu T^2}$ pour avoir quelques chances de déceler la sinusoïde à fréquence variable.

b. Autocorrélation.

1) $g(t) = k\, e^{2i\pi\nu(t+\varphi)}$:

$$E[\rho_g(h)] = \lim_{T\to\infty} \frac{1}{T} \int_0^T g(t)\, g(t+h)\, dt = \frac{1}{T} \int_0^T k^2 e^{2i\pi(t+\varphi)\nu} e^{-2i\pi\nu(t+h+\varphi)}\, dt = k^2 e^{-2i\pi\nu h}$$

Dans le cas où $g(t)$ est une sinusoïde réelle, on obtient :

$$\rho_g(h) = \frac{k^2}{2} \cos 2\pi h\nu$$

Lorsque $y(t)$ est la somme de réflexions, de bruits et d'une sinusoïde, l'autocorrélation de $y(t)$ est égale à la somme des autocorrélations des réflexions, de la sinusoïde, des bruits divers et des intercorrélations de ces divers facteurs qui tendent vers zéro quand $T \to \infty$.

2) $g(t) = k\, e^{2i\pi\nu(1+ct)t}$:

$$\rho_g(h) = \frac{k^2}{T} \int_0^T e^{2i\pi\nu[(1+ct)t - (1+c(t+h))(t+h)]}\, dt = \frac{k^2}{T} \int_0^T e^{-2i\pi\nu[(1+ch)h + 2cth]}\, dt$$

$$= -\frac{k^2 e^{-2i\pi\nu(1+ch)h + 2cth}}{4Ti\pi\nu ch} \left[e^{-4i\pi\nu cth}\right]_0^T = k^2 e^{-2i\pi\nu[(1+ch)h + cTh]} \frac{\sin 2\pi\nu cTh}{2\pi\nu cTh}$$

Lorsque $c = 0$, on retrouve $k^2 e^{-2i\pi\nu h}$.

Pour que la présence de la sinusoïde dans l'autocorrélation soit apparente, il est souhaitable que $\frac{\sin 2\pi\nu c\,T h}{2\pi\nu c\,T h} > \frac{1}{3}$; soit $2\pi\nu c\,T h < 2$. On obtient : $c < \frac{1}{\pi\nu T h}$

3) Influence d'un trend.

Il arrive que la droite des amplitudes moyennes d'une trace soit mal déterminée (1). En ce cas, un terme linéaire $m + lt$ avec l petit s'ajoute au signal théorique $y(t)$. On peut se ramener au cas où $m = 0$.

L'autocorrélation devient alors :

$$\rho(h) = \frac{1}{T}\int_0^T [lt + y(t)]\,[l(t+h) + y(t+h)]\,dt$$

$$\rho(h) = \frac{1}{T}\int_0^T y(t)\,y(t+h)\,dt + \frac{1}{T}\int_0^T [l(t+h)\,y(t) + y(t+h)\,lt]\,dt + \int_0^T \frac{l^2 t(t+h)}{T}\,dt$$

Le premier terme est inchangé, la troisième intégrale est égale à :

$$\frac{l^2 T^2}{3} + \frac{l^2 h T}{2}$$

Dans la seconde intégrale, l'intercorrélation du terme linéaire et des réflexions :

$$\int_{t-\varepsilon}^{t} s(t-\tau)\,k\left(\frac{\tau}{2}\right)d\tau$$

est nulle en espérance mathématique. Dans le cas où $y(t)$ comprendrait un terme en $k e^{2i\pi\nu t}$, on aurait :

$$\frac{kl}{T}\int_0^T \left[(t+h)\,e^{2i\pi\nu t} + t\,e^{-2i\pi\nu(t+h)}\right]dt$$

$$= \frac{kl}{\pi\nu}\left[e^{-2i\pi\nu\frac{h}{2}}\sin 2\pi\nu\left(T+\frac{h}{2}\right) + \frac{h}{T}e^{i\pi\nu T}\sin\pi\nu T + \frac{e^{-2i\pi\nu\frac{h}{2}}}{2T\pi\nu}\cos 2\pi\nu\left(T+\frac{h}{2}\right) + \frac{e^{-i\pi\nu h}}{2T\pi\nu}\cos\pi\nu h\right]$$

Lorsque T devient grand, le seul terme important dans l'expression précédente est :

$$\frac{kl}{\pi\nu}\,e^{-2i\pi\nu\frac{h}{2}}\sin 2\pi\nu\left(T+\frac{h}{2}\right)$$

En définitive, on n'identifie les divers termes de $\rho(h)$ que dans la mesure où $l < \frac{1}{T}$. Sinon, la présence du terme linéaire brouille complètement l'autocorrélation des réflexions et d'une éventuelle sinusoïde.

---

(1) En particulier lorsqu'on relève les amplitudes sur un film où la droite des amplitudes moyennes n'est pas tracée et doit être estimée.

c. Énergie relative des composantes de $y\ (t)$.

Si un signal $y\ (t)$ contient des réflexions, des termes sinusoïdaux et des bruits, l'autocorrélation, calculée sur un secteur T tendant vers l'infini donnera :

$$\rho\ (h) = \int_{-\mathfrak{T}}^{\mathfrak{T}} \rho_k \left(\frac{h-x}{2}\right) \rho_s\ (x)\ dx + \sum_i \frac{k_i^2}{2} \cos 2\pi\nu_i h + \mathrm{E}\ (h) + \eta\ (h)$$

$$\rho\ (0) = \int_{-\mathfrak{T}}^{\mathfrak{T}} \rho_k \left(\frac{-x}{2}\right) \rho_s\ (x)\ dx + \sum_i \frac{k_i^2}{2} + \mathrm{E}\ (0) + \eta\ (0)$$

E (h) est l'autocorrélation des bruits $\varepsilon\ (t)$ et $\eta\ (h)$ est l'intercorrélation des diverses composantes de $y\ (t)$.

Pour $h = 0$, et dans la mesure où T est assez long pour que $\eta\ (0)$ soit négligeable, $\rho\ (0)$ est la somme des énergies des réflexions, des sinusoïdes et du bruit.

Pour identifier la part d'énergie due à une sinusoïde, on doit relever les valeurs de $|\rho\ (h)|$ pour des valeurs de h multiples de la période et de la demi-période de la sinusoïde. Le rapport de leur moyenne à $\rho\ (0)$ indique la proportion d'énergie portée par la sinusoïde. On trouvera une illustration de cette méthode au chapitre VIII.

d. Remarques sur l'utilisation des autocorrélations et intercorrélations en sismique.

Pour obtenir des autocorrélations ou intercorrélations significatives, il faut effectuer les calculs sur un secteur T assez long, avec un échantillonnage assez serré des $y\ (t)$. Les divers calculs que nous avons effectués portaient sur des sommations faites sur ordinateur IBM 704. Les secteurs variaient de 500 à 800 ms, et l'échantillonnage des $y\ (t)$ était de 2 ou 4 ms, alors que les périodes les plus courtes sur $y\ (t)$ étaient d'environ 20 ms.

Il est dangereux de calculer les autocorrélations sur des secteurs beaucoup plus longs, car la période des impulsions varie au cours du temps et les conditions de stationnarité ne sont plus respectées. La stationnarité du second ordre est en outre très imparfaitement respectée, en dépit des contrôles de gain automatiques à l'enregistrement.

Nous avons vu que l'écart type de $\rho\ (h)$ normée sur des secteurs de 500 ou 800 ms est de l'ordre de 0,18 à 0,21. Ceci a pour conséquence qu'il est très difficile de déceler la présence d'une sinusoïde, même à période constante, ou d'un trend ; les larges oscillations de $\rho\ (h)$ pour $h > \mathfrak{T}$ traduisent plus souvent l'effet du hasard que celui de composantes périodiques. Cette variance élevée de l'autocorrélation enlève beaucoup d'intérêt à son emploi.

VI. CORRÉLATION MOBILE ET FILTRAGE. CONTROLE DE GAIN

1. Corrélation mobile.

Afin d'éliminer certaines fréquences indésirables, on peut procéder à un filtrage lors de l'enregistrement. Ce filtrage peut également être effectué à partir du signal enregistré sur le sismogramme.

Thèse J. AGARD. IV.

Le filtrage le plus simple, dans la mesure où $s\,(t)$ est connu, consiste à calculer la corrélation mobile de $y\,(t)$ par $s\,(t)$:

$$r\,(t) = \int_0^{\mathfrak{T}} s\,(\tau)\, y\,(t+\tau)\, d\tau$$

$$= \int_0^{\mathfrak{T}} \int_{t+\tau-\mathfrak{T}}^{t+\tau} s\,(\tau)\, s\,(t+\tau-\tau')\, k\left(\frac{\tau'}{2}\right) d\tau' d\tau + \int_0^{\mathfrak{T}} [g\,(t+\tau) + \varepsilon\,(t+\tau)]\, s\,(\tau)\, d\tau$$

Le domaine d'intégration de τ et τ' est celui de la zone hachurée de la figure 17. Avec un changement de variable $t - \tau' = u$, les intégrales en τ et u se séparent et l'on intègre en τ, puis en u. Les

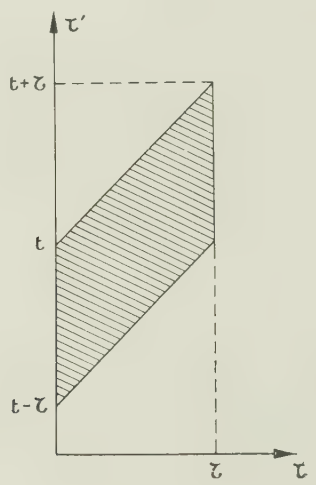


FIG. 17.

limites de l'intégrale en u dépendent de τ, mais elles peuvent être étendues à un intervalle plus large $(-\mathfrak{T}, +\mathfrak{T})$ car τ variant de 0 à $\mathfrak{T}$, u variera au plus de $-\mathfrak{T}$ à $\mathfrak{T}$, et il n'y a pas d'inconvénient à élargir le domaine de u $(-\tau < u < -\tau + \mathfrak{T})$ car $s\,(\tau + u)$ est nul hors de ce domaine. On obtient donc :

$$r\,(t) = \int_{-\mathfrak{T}}^{\mathfrak{T}} \rho_s\,(u)\, k\left(\frac{t-u}{2}\right) du + \int_0^{\mathfrak{T}} [g\,(t+\tau) + \varepsilon\,(t+\tau)]\, s\,(\tau)\, d\tau$$

Dans la mesure où les spectres de $g\,(t)$ et $\varepsilon\,(t)$ sont disjoints du spectre de $s\,(t)$, la corrélation mobile de $y\,(t)$ par $s\,(t)$ élimine presque complètement les bruits $g\,(t)$ et $\varepsilon\,(t)$. On en trouve une illustration au chapitre VIII.

La corrélation mobile peut du reste être considérée comme le résultat du passage du signal $y\,(t)$ dans un filtre R caractérisé par la fonction $s\,(-t)$ de gain $\Gamma\,(f)$:

$$\Gamma\,(f) = \int_{-\mathfrak{T}}^{0} s\,(-t)\, e^{-2i\pi ft}\, dt = \int_0^{\mathfrak{T}} s\,(t)\, e^{2i\pi ft}\, dt$$

En effet, appliquons le filtre R au signal $y\,(t)$:

$$\mathrm{R}\,[y\,(t)] = \int_{-\infty}^{+\infty} s\,[-(t-\theta)]\,y\,(\theta)\,d\theta$$

$$= \int_{t}^{t+\mathfrak{C}} s\,(\theta - t)\,y\,(\theta)\,d\theta \text{ car } s\,(\theta) = 0 \text{ si } t > \theta > t + \mathfrak{C}$$

$$= \int_{0}^{\mathfrak{C}} s\,(u)\,y\,(t+u)\,du \text{ avec } \theta - t = u$$

En passant n fois le signal $y\,(t)$ à travers le filtre R, on obtient un signal $\mathrm{R}_n\,(t)$ analogue au signal obtenu après le passage de $y\,(t)$ à travers un filtre de gain :

$$\Gamma_n\,(f) = [\Gamma\,(f)]^n$$

Lorsque $n \to \infty$, le spectre $\Gamma_n\,(f)$ tend vers un spectre de passe-raie :

$$\mathrm{R}_n\,(t) = \int_{0}^{\mathfrak{C}} \mathrm{K} y\,(t+\tau)\,e^{2i\pi f_0 \tau}\,d\tau$$

(f_0 est la fréquence où $\Gamma\,(f)$ à son module maximal).

La corrélation mobile a pour effet de déphaser le signal $y\,(t)$. En effet, un fort coefficient de réflexion en $\frac{\tau_0}{2}$ se traduit par un maximum de $y\,(t)$ en $\tau_0 + \mathrm{U}$ où U est l'abscisse du maximum de $s\,(t)$. Comme :

$$r\,(t) = \int_{-\mathfrak{C}}^{\mathfrak{C}} \rho_s\,(u)\,k\left(\frac{t-u}{2}\right)du = \int_{t-\mathfrak{C}}^{t+\mathfrak{C}} \rho_s\,(t-v)\,k\left(\frac{v}{2}\right)dv$$

on voit que $r\,(t)$ traduira la présence de $k\left(\frac{\tau_0}{2}\right)$ en $\tau_0 = t$.

Exemple d'un filtrage multiple.

Prenons le cas où $s\,(t) = s\,(-t) = e^{-it}$ pour $-\mathfrak{C} \leqslant t \leqslant \mathfrak{C}$ et $s\,(t) = 0$ si $-\mathfrak{C} > t > \mathfrak{C}$ et supposons $y\,(t)$ constitué d'un seul signal e^{-it} débutant en $t_1 - \mathfrak{C}$ et s'achevant en $t_1 + \mathfrak{C}$. La corrélation mobile sera :

$$r\,(t_1) = \frac{1}{2\mathfrak{C}} \int_{-\mathfrak{C}}^{\mathfrak{C}} e^{-it}\,e^{-i(t-t_1)}\,dt$$

Le gain du filtre de caractéristique e^{it} sera :

$$\Gamma\,(f) = \frac{1}{2\mathfrak{C}} \int_{-\mathfrak{C}}^{\mathfrak{C}} e^{it}\,e^{-2i\pi ft}\,dt = \frac{\sin \mathfrak{C}\,(1-2\pi f)}{\mathfrak{C}\,(1-2\pi f)} < 1$$

Après n passages de $y\,(t)$ dans le filtre, on obtiendra un signal $r_n\,(t)$ analogue au passage de $y\,(t)$ dans un filtre de gain $\Gamma_n\,(f) = \left[\frac{\sin \mathfrak{C}\,(1-2\pi f)}{\mathfrak{C}\,(1-2\pi f)}\right]^n$. Ce filtre tend vers 0 quand $n \to \infty$, sauf en $f = \frac{1}{2\pi}$

où $\Gamma_n\left(\frac{1}{2\pi}\right) = 1$. Lorsque $n \to \infty$, le filtrage de $y(t)$ par $\Gamma_n(f)$ donne un signal $r_n(t)$ qui tend vers zéro pour tout t.

Remarque sur la corrélation mobile.

On peut adopter diverses formes pour la corrélation mobile. La forme $r(t)$ adoptée plus haut est la plus simple. On peut adopter une forme plus voisine de celle de l'autocorrélation normée avec :

$$r(t) = \frac{\int_0^{\mathfrak{T}} s(\tau)\, y(t+\tau)\, d\tau}{\sqrt{\int_0^{\mathfrak{T}} y^2(t+\tau)\, d\tau}}$$

Cette forme fournit un signal moins bien corrélé avec $y(t)$ car les fortes réflexions y sont amoindries par la normalisation. Elle nous semble cependant apporter une information complémentaire dans la mesure où $s(t)$ est bien connu, car un maximum de $r(t)$ indique le départ d'une réflexion qui sera d'autant plus certaine que $r(t)$ sera grand. Le maximum correspondant de $y(t)$ indique l'intensité de la réflexion et $r(t)$ la qualité de la forme de la réflexion par rapport à $s(t)$.

En fait, les courbes de $r(t)$ normées ou non sont assez voisines, et une réflexion étant souvent constituée par la somme de deux ou trois coefficients de réflexion voisins, il faudrait corréler non par $s(t)$ mais par :

$$\int_0^u s(t+\tau)\, k\left(\frac{\tau}{2}\right) d\tau$$

expression inconnue a priori.

2. Contrôle de gain.

La méthode de la corrélation mobile suppose la connaissance de $s(t)$. A priori $s(t)$ est inconnu. Pour appliquer la méthode de corrélation mobile, certains (10) ont pensé à émettre un signal $s(t)$ connu et très long par la rotation accélérée, puis freinée d'un volant à balourd, dont le bâti est couplé au sol. Nous donnons dans la figure 18 la formule d'un tel signal avec le dessin de son autocorrélation qui est assez raide à l'origine pour que la corrélation mobile fasse bien ressortir les coefficients de réflexion élevés.

La méthode nous semble cependant sujette à plusieurs critiques. Le signal émis dure quatre secondes. De ce fait, reviennent en surface en même temps la fin du signal $s(t)$ provenant des couches superficielles et le début du signal revenant des couches les plus profondes et dont les intensités ne sont que le 1/100 des premières du fait de l'absorption.

L'emploi d'un contrôle de gain destiné à maintenir l'énergie reçue en surface à un niveau constant conservera ces inégalités relatives et les réflexions profondes risquent fort d'être étouffées.

a. Contrôle de gain et corrélation mobile.

En outre, la méthode de la corrélation mobile perd de son efficacité du fait du contrôle de gain. Soit $\gamma(t)$ le gain, supposé connu. Le signal corrigé est $y(t) \cdot \gamma(t)$. La corrélation mobile est :

$$r(t) = \int_0^{\mathfrak{T}} s(u)\, \gamma(t+u)\, y(t+u)\, du$$

$$= \int_0^{\mathfrak{T}} s(u)\, \gamma(t+u) \left[\int_{t+u-\mathfrak{T}}^{t+u} s(t+u-\tau)\, k\left(\frac{\tau}{2}\right) d\tau + \varepsilon(t+u) + g(t+u) \right] du$$

Les deux derniers termes du crochet seront bien éliminés par le filtrage, mais le premier terme ne donne plus l'expression obtenue précédemment pour la corrélation mobile :

$$r(t) = \int_{-\mathfrak{T}}^{\mathfrak{T}} \rho_s(u)\, k\left(\frac{t-u}{2}\right) du$$

Même en corrélant $s(u)\,\gamma(u)$ avec $\gamma(t+u)\,y(t+u)$, on ne fait pas apparaître une expression du type précédent. On a :

$$\int_{o}^{\mathfrak{T}} \int_{t+u-\mathfrak{T}}^{t+u} s(u)\,\gamma(u)\,s(t+u-\tau)\,\gamma(t+u)\,k\left(\frac{\tau}{2}\right) d\tau\, du$$

On ne peut séparer l'intégrale donnant $\rho_s(u)$ ou $\rho_{s\gamma}(u)$. Cette perturbation due au contrôle de gain est très grave sur un signal long, mais insignifiante sur un signal $s(t)$ court car la variation de $\gamma(t)$ sur 60 ou 80 ms est très faible.

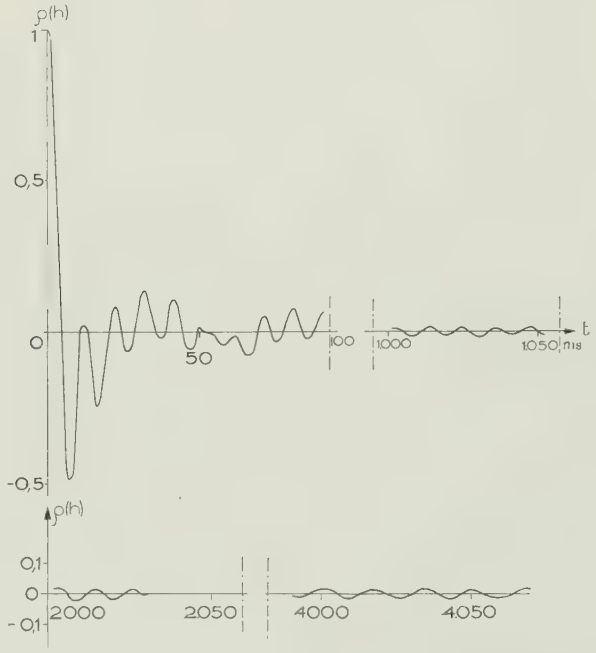


FIG. 18.
Autocorrélation de S(t).
Signal long.

$$S(t) = \begin{cases} y_1(t) = \sin 20\pi\,(t+2)\,t \\ \text{pour } o \geqslant t \geqslant 4 \text{ s.} \\ y_2(t) = \sin 20\pi\,(18-t)\,t \\ \text{pour } 4 \geqslant t \geqslant 8 \text{ s.} \end{cases}$$

b. Contrôle de gain et analyse harmonique.

$$Y(f) = \int_{o}^{T} y(t)\,\gamma(t)\,e^{-2i\pi ft}\,dt$$

$$= \int_{o}^{T} \gamma(t) \int_{t-\mathfrak{T}}^{t} s(t-\tau)\,k\left(\frac{\tau}{2}\right) e^{-2i\pi f(t-\tau)-2i\pi f\tau}\,d\tau\,dt + \int_{o}^{T} [g(t)+\varepsilon(t)]\,\gamma(t)\,e^{-2i\pi ft}\,dt$$

$$= \int_{o}^{T} \int_{o}^{\mathfrak{T}} \gamma(t)\,s(u)\,e^{-2i\pi fu}\,k\left(\frac{t-u}{2}\right) e^{-2i\pi f(t-u)}\,du\,dt + \Gamma G(f) + \Gamma E(f) \quad \text{avec} \quad u = t-\tau$$

où ΓG et ΓE sont les spectres de $g(t)\cdot\gamma(t)$ et de $\varepsilon(t)\cdot\gamma(t)$.

$$Y(f) = \int_0^{T-u} \int_0^{\mathfrak{G}} \gamma(u+v)\, s(u)\, e^{-2i\pi fu}\, k\left(\frac{v}{2}\right) e^{-2i\pi fv}\, du\, dv + \Gamma G(f) + \Gamma E(f)$$

L'introduction du gain interdit la mise en évidence du produit $S(f) \cdot K(f)$ que nous obtenions au chapitre V.

c. Contrôle de gain et autocorrélation.

$$\rho(h) = \frac{1}{T} \int_0^T \gamma(t)\, y(t)\, \gamma(t+h)\, y(t+h)\, dt$$

$$= \frac{1}{T} \int_0^T \int_{t-\mathfrak{G}}^{t} \gamma(t)\, s(t-\tau)\, k\left(\frac{\tau}{2}\right) d\tau \int_{t+h-\mathfrak{G}}^{t+h} \gamma(t+h)\, s(t+h-\tau')\, k\left(\frac{\tau'}{2}\right) d\tau'\, dt$$

$$+ \frac{1}{T} \int_0^T \gamma(t)\, \gamma(t+h)\, [g(t+h) + \varepsilon(t+h)] \int_{t-\mathfrak{G}}^{t} s(t-\tau)\, k\left(\frac{\tau}{2}\right) d\tau\, dt$$

$$+ \frac{1}{T} \int_0^T \gamma(t)\, \gamma(t+h)\, [g(t) + \varepsilon(t)] \int_{t+h-\mathfrak{G}}^{t+h} s(t+h-\tau')\, k\left(\frac{\tau'}{2}\right) d\tau'\, dt$$

$$+ \frac{1}{T} \int_0^T [g(t) + \varepsilon(t)]\, [g(t+h) + \varepsilon(t+h)]\, \gamma(t)\, \gamma(t+h)\, dt$$

$$= \frac{1}{T} \int_0^T \int_{t-\mathfrak{G}}^{t} \int_{t+h-\mathfrak{G}}^{t+h} s(t-\tau)\, s(t+h-\tau')\, \gamma(t)\, \gamma(t+h)\, k\left(\frac{\tau}{2}\right) k\left(\frac{\tau'}{2}\right) d\tau\, d\tau'\, dt$$

$$+ \frac{1}{T} \int_0^T g(t)\, g(t+h)\, \gamma(t)\, \gamma(t+h)\, dt + \eta(h)$$

$\eta(h)$ étant constitué des résidus d'intégrales nulles en moyenne. On ne peut faire apparaître dans la première intégrale de terme en $\rho_{\gamma s}(x)\, \rho_k\left(\frac{h-x}{2}\right)$ comme précédemment.

d. Influence sur $y(t)$.

Le contrôle de gain ayant pour but de rendre le signal stationnaire, l'utilisation d'un contrôle automatique est dangereuse, car elle a pour résultat de réduire les grandes amplitudes et de majorer les faibles, c'est-à-dire de réduire l'amplitude des réflexions notables et d'augmenter la probabilité des fausses réflexions.

VII. DÉTERMINATION DE S (T) PAR L'AUTOCORRÉLATION

Dans la mesure où les coefficients de réflexion sont répartis à la Poisson, l'autocorrélation de $y(t)$, signal d'une trace ne comportant pas de bruit, est identique à l'autocorrélation de l'impulsion $s(t)$. Certains auteurs, ROBINSON, HORTON, proposent d'utiliser cette propriété pour identifier $s(t)$.

1. Méthode de Robinson.

Dans les hypothèses de stationnarité, d'absence de bruit autocorrélé et de répartition des coefficients de réflexion à la Poisson, ROBINSON (25) expose les méthodes classiques de prévision linéaire. Il opère sur la transformée de Fourier de l'autocorrélation, soit : $\Phi(f) = |S(f)|^2$. Il préfère cette méthode à celle de l'analyse harmonique, parce qu'il indique que l'analyse harmonique de $y(t)$ donnerait :

$$Y(f) = S(f) \cdot K(f).$$

$K(f)$: spectre des $k(\tau)$.

La connaissance de $Y(f)$ calculé à partir de $y(t)$ ne lui apprendrait rien sur $s(t)$, alors qu'avec les hypothèses de répartition à la Poisson il obtient une autocorrélation $\rho(h) = C\,\rho_s(h)$. Il nous semble que ce raisonnement contienne une contradiction car, dans la mesure où l'on suppose les coefficients de réflexion répartis à la Poisson, ils ont une autocorrélation nulle sauf en $h = 0$, leur spectre est constant ($K(f) = K$) et l'analyse harmonique de $y(t)$ fournit directement le spectre de $s(t)$. Dans la mesure où le secteur T sur lequel on effectue les calculs ne tend pas vers l'infini, ces résultats sont seulement approximatifs et la variance sur l'autocorrélation ou le spectre théorique des $k(\tau)$ est importante. On n'a donc pas :

$$K(f) = K \quad \text{ou} \quad \rho_k(h) = 0 \quad \text{pour} \quad h \neq 0$$

Estimant connaître $|S(f)|^2$ par l'intermédiaire de l'autocorrélation, ROBINSON cherche l'opérateur inverse de $s(t)$, dont le spectre $= \frac{1}{S(f)}$. En appliquant cet opérateur sur $y(t)$, il obtient en théorie un signal dont le spectre est :

$$K(f) \cdot S(f) \cdot \frac{1}{S(f)} = K(f)$$

La détermination de l'opérateur inverse est obtenue en supposant que $S(\lambda + i\mu)$ n'a pas de singularité dans le demi-plan imaginaire $\mu > 0$ ($f = \lambda + i\mu$), c'est-à-dire que $s(t) = 0$ si $t < 0$. L'opérateur est stable et a une caractéristique de phase minimale. Cette méthode dont le détail des calculs pratiques n'est pas indiqué, paraît inapplicable dans les conditions réelles.

2. Méthode de Horton.

Dans un article de *Geophysics* (17), HORTON propose une méthode de détermination de l'impulsion $s(t)$ que nous avons testée dans divers cas simples avec des succès inégaux. Il utilise la méthode suivante : il suppose que le signal $s(t)$ est donné par une série de Gram Charlier :

$$s(t) = \sum_{n=0}^{\infty} b_n \Phi^{(n)}\left(\frac{t}{\mathfrak{T}}\right), \qquad \left(\Phi^{(n)}(x) = \frac{d^n}{dx^n}\,\frac{\left(e^{\frac{-x^2}{2}}\right)}{\sqrt{2\pi}}, \; x = \frac{t}{\mathfrak{T}}\right)$$

Une telle forme de signal $s(t)$ correspond à certaines hypothèses de propagation des ondes de Ricker. Les signaux sont assez représentatifs de ce que nous observons sur des sismogrammes.

Les termes $\Phi^{(n)}(x)$ étant orthogonaux, l'autocorrélation d'un signal $s(t)$ est de la forme :

$$\rho_s\left(\frac{\tau}{\mathfrak{T}}\right) = \sum_{(m,n)=1}^{\infty} (-1)^m a_m a_n \Phi^{(m+n)}\left(\frac{\tau}{\mathfrak{T}\sqrt{2}}\right), \qquad \left(a_m = \left(\frac{2^{-m} N \mathfrak{T}}{\sqrt{2}}\right)^{\frac{1}{2}} b_m\right)$$

N est le nombre de réflexions par unité de temps.

Comme nous connaissons en réalité l'autocorrélation empirique, mais que nous ignorons les valeurs des b_m du signal, il faut déterminer les b_m à partir de l'autocorrélation empirique.

Écrivons cette autocorrélation :

$$\sum_{p=1} c_p \Phi^{(p)}\left(\frac{y}{\sqrt{2}}\right) = \sum_{(m,n)=1}^{\infty} (-1)^m a_m a_n \Phi^{(m+n)}\left(\frac{y}{\sqrt{2}}\right)$$

On aura :

$$\begin{aligned}
c_2 &= -a_1^2 \\
c_4 &= -2a_1a_3 + a_2^2 \\
c_6 &= -2a_1a_5 + 2a_2a_4 - a_3^2 \\
c_{2m} &= -2a_1a_{2m} + 2a_2a_{2m-2} - 2a_3a_{2m-3} \\
&\quad \ldots + 2(-1)^{m-1}a_{m-1}a_{m+1} + (-1)^m a_{2m}^2 \\
c_{2m+1} &= 0 \text{ quel que soit M.}
\end{aligned}$$

L'ensemble précédent des équations n'est pas linéaire en a_i, de sorte que la solution n'est pas unique. Si l'on ne fait intervenir que trois ou quatre termes b_i, on peut étudier toutes les solutions possibles et obtenir une assez grande variété de signaux $s(t)$.

Pour déterminer les divers c_i, HORTON commence par annuler les valeurs de $\rho(t)$ au-delà d'une certaine valeur où il estime que $\rho(t)$ est nul en espérance mathématique. La détermination de cette valeur est arbitraire. Remarquons que l'on échappe difficilement à ce problème de la troncature. Si l'on passe de l'autocorrélation au périodogramme de SCHUSTER, on est ensuite amené à déterminer un point de troncature pour le périodogramme (29).

D'autre part, la détermination de $\mathfrak{T}$ destiné à normer t pour avoir une variable y sans dimension est également arbitraire et entraîne une hypothèse sur la longueur $\mathfrak{T}$ de l'impulsion $s(t)$.

Ayant obtenu une autocorrélation estimée $\hat{\rho}(y)$, HORTON calcule les coefficients c_{2m} avec :

$$c_{2m} = \frac{1}{(2m)!} \cdot \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} \hat{\rho}(y)\, H_{2m}(y\sqrt{2})\, dy,$$

H_{2m} : polynome d'Hermite.

On calcule ainsi tous les c_{2m}, en tenant compte de ce que $\sum_{m=0}^{\infty} c_{2m} \Phi^{(2m)}(0) = s \to 1$, ce qui permet de s'arrêter lorsqu'on approche de 1. On peut aussi comparer $\sum_m c_{2m} \Phi^{(2m)}\left(\frac{y}{\sqrt{2}}\right)$ avec $\hat{\rho}(t)$ pour vérifier la concordance de l'ajustement. En général, trois ou quatre termes donnent une approximation assez correcte de $\rho(t)$.

La méthode précédente peut être appliquée jusqu'au bout, mais elle offre plusieurs inconvénients graves : elle suppose une répartition rigoureuse des coefficients de réflexion à la Poisson et elle suppose l'absence de tout bruit. Sinon, l'autocorrélation de $y(t)$ ne s'identifie plus à celle du signal. La forme du signal est déterminée a priori (série de Gram Charlier), mais cette forme est assez souple pour bien approcher un très grand nombre de signaux $s(t)$. Par contre, plus on introduit de termes b_n, plus on complique la résolution qui offre un certain nombre d'indéterminations.

Enfin, nous avons vu au chapitre V, paragraphe 4 que l'autocorrélation présente une variance très importante. Sur des sismogrammes synthétiques dépourvus de tout bruit, calculés à l'*Institut Français du Pétrole*, nous avons obtenu des autocorrélations très différentes de celle de l'impul-

sion $s(t)$. La troncature de l'autocorrélation de la trace se révèle donc très dangereuse et elle risque de fausser complètement les calculs ultérieurs.

3. Assimilation de $s(t)$ à une période de sinusoïde.

La méthode de HORTON nécessite un grand nombre de calculs entraînant souvent des approximations qui font perdre beaucoup d'intérêt à la précision possible théoriquement.

D'après certains résultats, il semble qu'une simple période de sinusoïde suffise à représenter le signal $s(t)$. En ce cas, l'autocorrélation théorique du signal serait, pour un signal :

$$s(x) = \begin{cases} \sin \beta x \ldots\ldots\ldots\ldots\ldots\ldots & \text{si} \quad 0 \leqslant x \leqslant \dfrac{2\pi}{\beta} \\ 0. \ldots\ldots\ldots\ldots\ldots\ldots & \text{ailleurs} \end{cases}$$

$$\rho(t) = \int_0^{2\frac{\pi}{\beta}-t} \sin \beta x \sin \beta (x+t)\, dx = \frac{1}{2} \int_0^{2\frac{\pi}{\beta}-t} [\cos \beta t - \cos \beta (2x+t)]\, dx$$

$$= \frac{1}{2}\left[\left(\frac{2\pi}{\beta} - t\right)\cos \beta t + \frac{\sin \beta t}{\beta}\right]$$

$$\frac{\partial \rho(t)}{\partial t} = -\frac{1}{2}\left(\frac{2\pi}{\beta} - t\right)\beta \sin \beta t$$

$$\frac{\partial \rho(t)}{\partial t} = 0 \text{ pour } t = 0, \qquad t = \frac{\pi}{\beta} \qquad \text{et} \qquad t = \frac{2\pi}{\beta}$$

$$\rho(0) = \frac{\pi}{\beta}, \qquad \rho\left(\frac{\pi}{\beta}\right) = \frac{-\pi}{2\beta}, \qquad \rho\left(\frac{2\pi}{\beta}\right) = 0$$

Une simple période de sinusoïde pour représenter le signal ne semble pas suffisante car nous aurions alors toujours pour premier minimum de $\rho(t)$ normé (fig. 19) la valeur $-\dfrac{\rho(0)}{2}$ en $t = \dfrac{\mathfrak{T}}{2}$

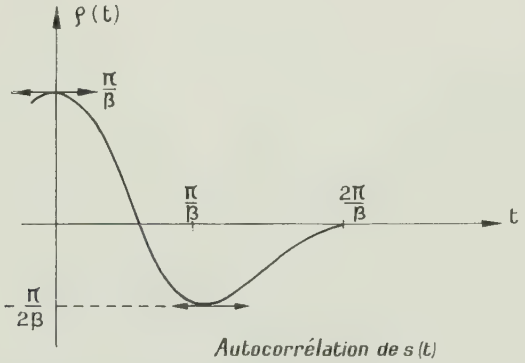


FIG. 19. *Autocorrélation de s(t)*

($\mathfrak{T}$ étant la période du signal). Dans la plupart des cas pratiques, il ne semble pas que ce soit le cas Ce schéma très simple d'un ajustement immédiat, connaissant la position du premier minimum de $\rho(\tau)$, ne semble pas convenable.

4. Essai d'une sinusoïde amortie $y(x) = e^{-\alpha x} \sin \beta x$.

$$\rho(t) = \int_0^\infty e^{-2\alpha x - \alpha t} \sin \beta x \sin \beta (x + t)\, dx$$

$$= \frac{1}{2} \int_0^\infty e^{-\alpha(2x+t)} [\cos \beta t - \cos \beta (2x + t)]\, dx = \frac{e^{-\alpha t} \cos \beta t}{2} \int_0^\infty e^{-2\alpha x}\, dx$$

$$- \frac{1}{2} \int_0^\infty e^{-\alpha(2x+t)} \cos \beta (2x + t)\, dx$$

Par un changement $y = 2x + t$, on obtient :

$$\rho(t) = \frac{e^{-\alpha t} \cos \beta t}{4\alpha} - \int_t^\infty \frac{e^{-\alpha y}}{4} \cos \beta y dy = \frac{e^{-\alpha t} \cos \beta t}{4\alpha} - \frac{A}{4}$$

$A = \int_t^\infty e^{-\alpha y} \cos \beta y dy$ se calcule avec $B = \int_t^\infty e^{-\alpha y} \sin \beta y dy$ par intégration par partie.

$$A = \frac{e^{-\alpha t}}{\alpha^2 + \beta^2} [\beta \sin \beta t + \alpha \cos \beta t]$$

Donc : $$\rho(t) = \frac{e^{-\alpha t}}{4\alpha(\alpha^2 + \beta^2)} [(\alpha^2 + \beta^2) \cos \beta t - \alpha [\beta \sin \beta t + \alpha \cos \beta t]]$$

$$\boxed{\rho(t) = \frac{\beta e^{-\alpha t}}{4\alpha(\alpha^2 + \beta^2)} (\beta \cos \beta t - \alpha \sin \beta t)}$$

$$\frac{d\rho(t)}{dt} = - \frac{\beta e^{-\alpha t} \sin \beta t}{4\alpha}$$

Les extremums de $\rho(t)$ sont situés en $t = \frac{k\pi}{\beta}$ (k entier positif ou nul).

$$\rho(0) = \frac{\beta^2}{4\alpha(\alpha^2 + \beta^2)}$$

$$\rho\left(\frac{\pi}{\beta}\right) = \frac{-\beta^2 e^{-\frac{\alpha\pi}{\beta}}}{4\alpha(\alpha^2 + \beta^2)}$$

. . . . . .

$$\rho\left(\frac{k\pi}{\beta}\right) = \frac{\beta^2 e^{-k\frac{\alpha\pi}{\beta}}}{4\alpha(\alpha^2 + \beta^2)} \cdot (-1)^k$$

Les valeurs successives des extremums se déduisent de $\rho(0)$ en multipliant $\rho(0)$ par $(-1)^k e^{-\frac{\alpha\pi}{\beta} k}$ ($k = 0, 1, \ldots$)

On voit que β est déterminé par la position des extremums, et α par les valeurs de ces extremums.

L'assimilation du signal $s(t)$ à un signal de la forme $e^{-\alpha t} \sin \beta t$ permet la détermination des constantes α et β par la seule connaissance de l'autocorrélation de $y(t)$, dans le cas où les $k(\tau)$ sont répartis à la Poisson.

5. Détermination d'un signal de Ricker.

L'impulsion $s(t)$ est souvent assimilée à un signal de Ricker :

$$s(t) = \frac{\sqrt{\pi}}{m}\left[1 - 2\left(\frac{t}{m}\right)^2\right] e^{-\left(\frac{t}{m}\right)^2}$$

L'autocorrélation théorique de cette impulsion s'obtient par des intégrations par partie et est égale à :

$$\rho_s(h) = \frac{h^4 - 6h^2m^2 + 3m^4}{16\pi m^5}\sqrt{\frac{\pi}{2}}\, e^{-\frac{h^2}{2m^2}}$$

Le signal de Ricker contient un seul paramètre : m. Il suffit de connaître une valeur de $\rho(h)$ avec $h \neq 0$ pour en tirer la valeur de m. En particulier, on peut voir que l'autocorrélation s'annule en $h = \pm m\sqrt{3 \pm \sqrt{6}}$ et qu'elle atteint son minimum en $h = \pm m\sqrt{5 - \sqrt{10}}$ et ses maximums en $h = 0$ et $h = \pm m\sqrt{5 + \sqrt{10}}$.

On peut calculer les valeurs de m correspondant aux valeurs de h lorsque l'autocorrélation s'annule ou passe par ses extremums et adopter pour m leur moyenne.

Remarquons que les zéros et les extremums de $s(t)$ ne coïncident pas avec ceux de $\rho_s(h)$, car $s(t)$ s'annule en $\pm \frac{m}{\sqrt{2}}$ et a ses extremums en $t = 0$ et $t = \pm \frac{m\sqrt{6}}{2}$.

Enfin, l'autocorrélation $\rho_s(h)$ étant égale à $\frac{\sqrt{\pi}}{m}$ en $h = 0$ et à $-\frac{\sqrt{\pi}}{m} 2e^{-\frac{3}{2}}$ pour $h = \pm m\sqrt{5 - \sqrt{10}}$, on voit que l'autocorrélation normée est égale à $-2e^{-\frac{3}{2}}$ en $h = \pm m\sqrt{5 - \sqrt{10}}$.

6. Méthode de Lindsey : élimination des réflexions fantômes.

Dans un article de *Geophysics* (18), LINDSEY indique une méthode destinée à éliminer les « réflexions fantômes » dues aux tirs profonds.

On voit sur la figure 20 que l'impulsion de l'explosion peut suivre le chemin 1, ce qui la fait revenir en surface à l'instant t après réflexion sur la couche n_0, ou elle peut suivre le chemin 2 qui la ramène en surface au temps $t + \tau$ après réflexion en surface, puis sur la couche n_0. On montre très facilement que, si a est le coefficient de réflexion de la surface, la transformée de Laplace du signal reçu en surface sera :

$$K(s)\left[S(s) - aH(s)\, S(s)\, e^{-s\tau}\right]$$

$K(s)$: transformée de Laplace de la réponse du sol,
$S(s)$: transformée de Laplace de l'impulsion,
$H(s)$: transformée de Laplace traduisant l'effet du sol durant l'aller-retour entre le point de l'explosion et la surface.

L'utilisation d'un filtre caractérisé par :

$$F(s) = \frac{1}{1 - a\,H(s)\,e^{-s\tau}}$$

permet d'éliminer la première réflexion fantôme, qui est la plus importante, par un système à *feed-back*. L'estimation des paramètres a et τ peut être obtenue par l'autocorrélation. Comme la présence des réflexions fantômes transforme le signal $y(t)$ en $y(t) - a\,y(t-\tau)$, l'autocorrélation du signal reçu sera :

$$\rho(h) = (1 + a^2)\,\rho_y(h) - a\,\rho_y(h+\tau) - a\,\rho_y(h-\tau)$$

$\rho_y(h)$ est l'autocorrélation de $y(t)$ (signal sans réflexion fantôme). La présence des réflexions fantômes se traduit par la présence en $\pm\tau$ d'un minimum d'amplitude relative $m = \frac{a}{a^2+1}$ dans l'autocorrélation normée (si $\tau > \mathfrak{T}$, longueur de l'impulsion).

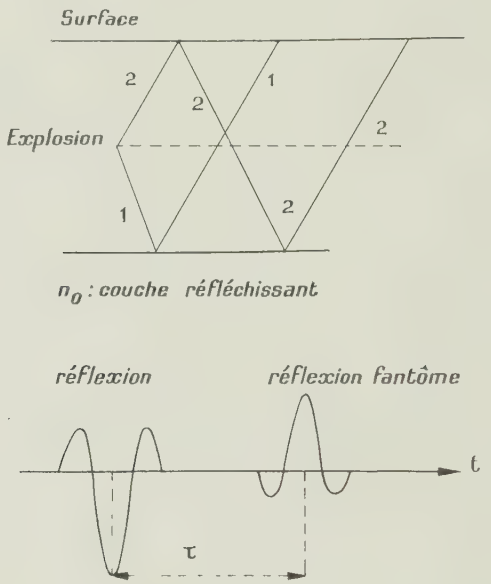


FIG. 20.

On suppose en fait que $H(s) = 1$. La détermination de m et de son emplacement dans l'autocorrélation normée, ainsi que la similitude de $\rho(h)$ entre $-\mathfrak{T}$ et $+\mathfrak{T}$ et de $-\rho(h+\tau)$ et $-\rho(h-\tau)$ entre $-\mathfrak{T}+\tau$ et $\tau+\mathfrak{T}$ et $-\mathfrak{T}-\tau$ et $-\tau+\mathfrak{T}$ permettent de déceler la présence des réflexions fantômes et de construire un filtre.

Cette méthode nécessite les mêmes hypothèses de stationnarité du signal et de répartition des coefficients de réflexion à la Poisson, que les méthodes de ROBINSON et HORTON. Dans les cas où ces hypothèses sont mal vérifiées, ou si les bruits sont importants, cette méthode nous semble peu applicable. La variance de l'autocorrélation risque de cacher l'autocorrélation des fantômes ou de suggérer leur présence là où elle n'est pas. Le délai τ ne correspond pas exactement au temps d'aller retour entre l'explosion et la surface, de sorte qu'on ne peut pas connaître à l'avance la position du maximum de l'autocorrélation des fantômes. La fonction $H(s)$ n'est généralement pas égale à 1. Cependant, dans la mesure où plusieurs raisons nous permettent de penser que des réflexions fantômes existent, la méthode est applicable.

Le filtrage proposé sera satisfaisant, mais il faut s'attendre que les bruits de surface soient multipliés puisque les bruits n'ont pas de « bruit fantôme » correspondant.

VIII. MÉTHODE DES SISMOGRAMMES SYNTHÉTIQUES

Afin d'essayer les différentes méthodes exposées ci-dessus et pour étudier l'influence des réflexions multiples, de bruits divers ou de fonctions périodiques sur un sismogramme, il est intéressant de construire des « sismogrammes synthétiques ». Plusieurs procédés peuvent être envisagés, soit le procédé analogique, avec la barre (8), soit avec les calculateurs électroniques, en utilisant des coeffi-

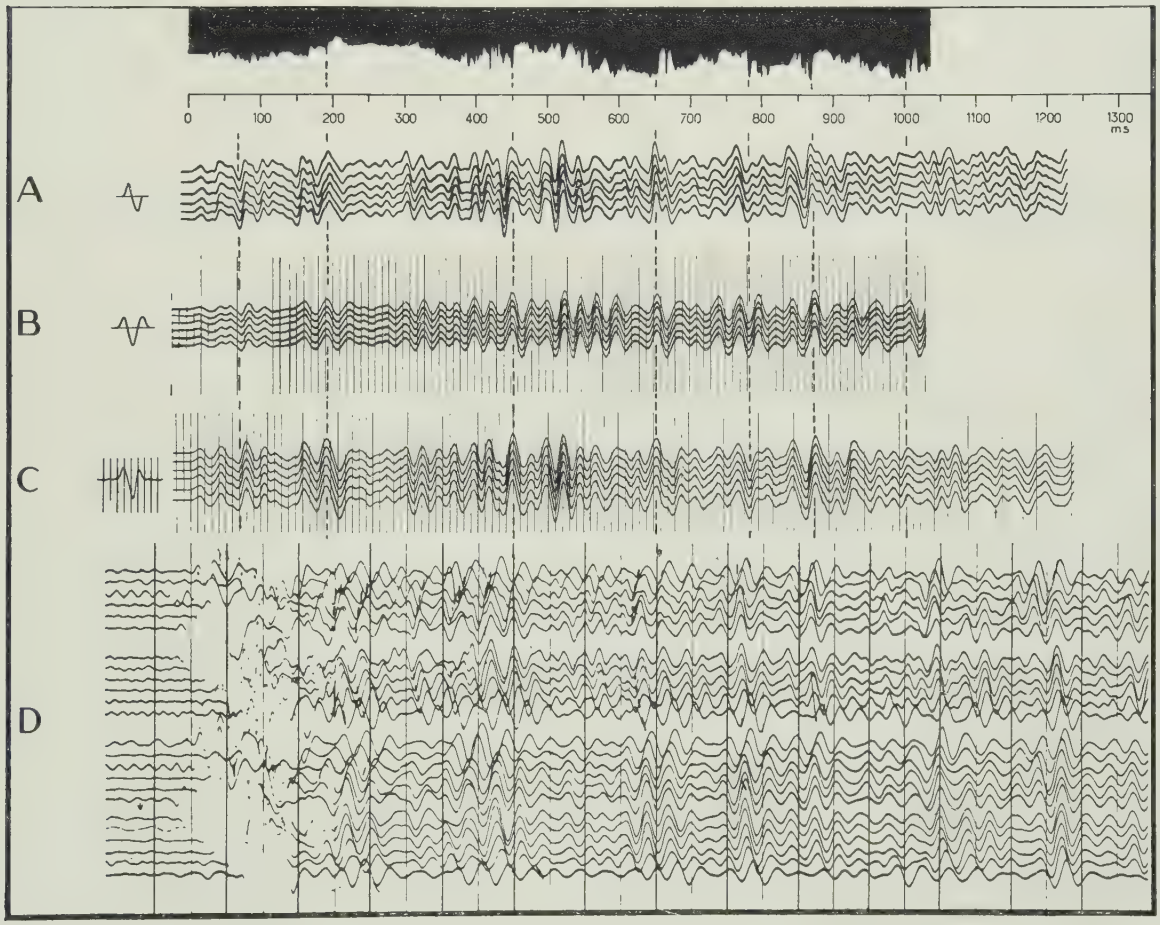


FIG. 21.
Comparaison de sismogrammes synthétiques à multiples et sans multiple avec le sismogramme.

en haut : courbe des vitesses.
A : sismogramme synthétique à multiples et son impulsion.
B : sismogramme synthétique sans multiple et son impulsion.
C : sismogramme synthétique à multiples et son impulsion.
D : sismogramme réel { en haut, non mixé ; en bas, mixé.

cients de réflexion provenant d'un carottage de vitesses, soit en fabriquant des coefficients de réflexion synthétiques par des tirages au hasard reproduisant les conditions aléatoires réelles.

Nous donnons ici quelques exemples de sismogrammes synthétiques. La figure 21 permet de faire la comparaison entre des sismogrammes synthétiques à multiples et sans multiple et un sismogramme réel correspondant.

Les sismogrammes synthétiques à multiples peuvent être calculés par la méthode de GRAU (2) et (8). Ils indiquent un accord satisfaisant avec les sismogrammes synthétiques sans multiple durant les premières 500 ou 600 ms du film, puis les différences deviennent plus grandes.

1. Calcul du sismogramme.

Les sismogrammes synthétiques sans multiple sont calculés à partir de la donnée de l'impulsion $s(t)$ et des coefficients de réflexion $k(t_j)$ échantillonnés toutes les 1 ou 2 ms selon les cas. On écrit :

$$y(t_i) = \sum_{j=i-m}^{i} s(t_i - t_j)\, k\left(\frac{t_j}{2}\right)$$

Nous avons ainsi utilisé les coefficients $k(t)$ du graphique (fig. 2) [1] pour fabriquer un sismogramme synthétique sans réflexion multiple. La figure 22 montre les sismogrammes synthétiques correspon-

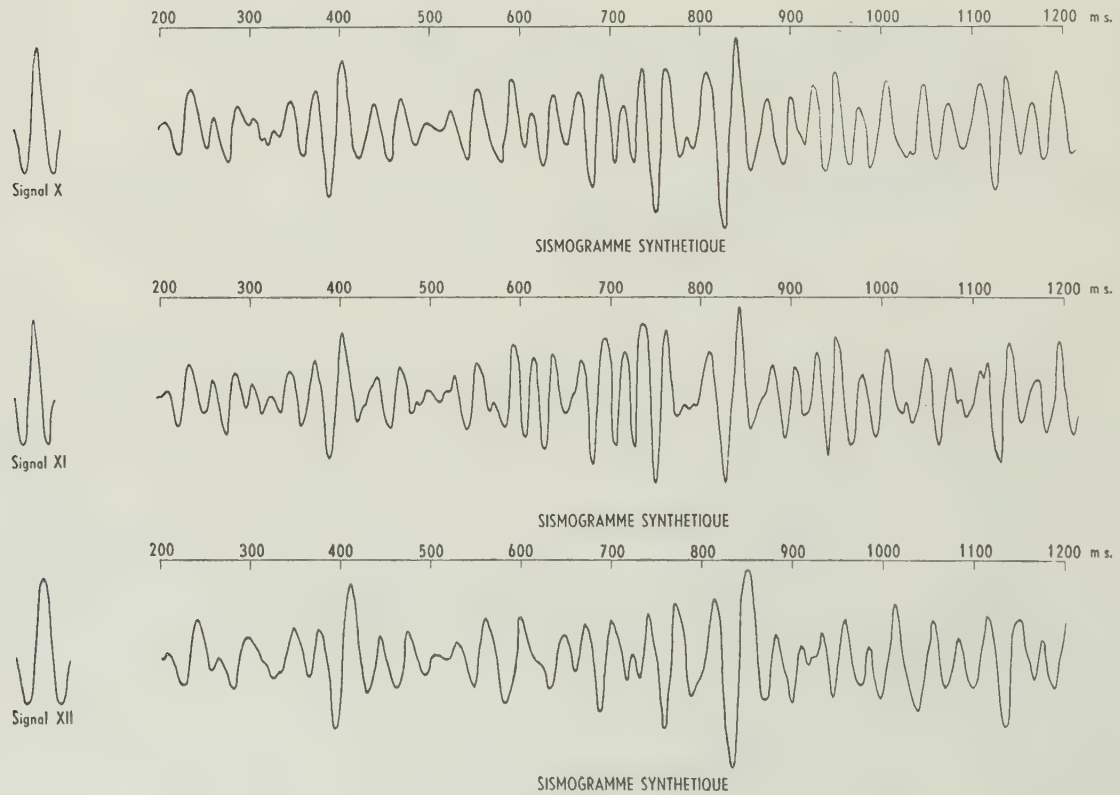


FIG. 22.

[1] Voir p. 9.

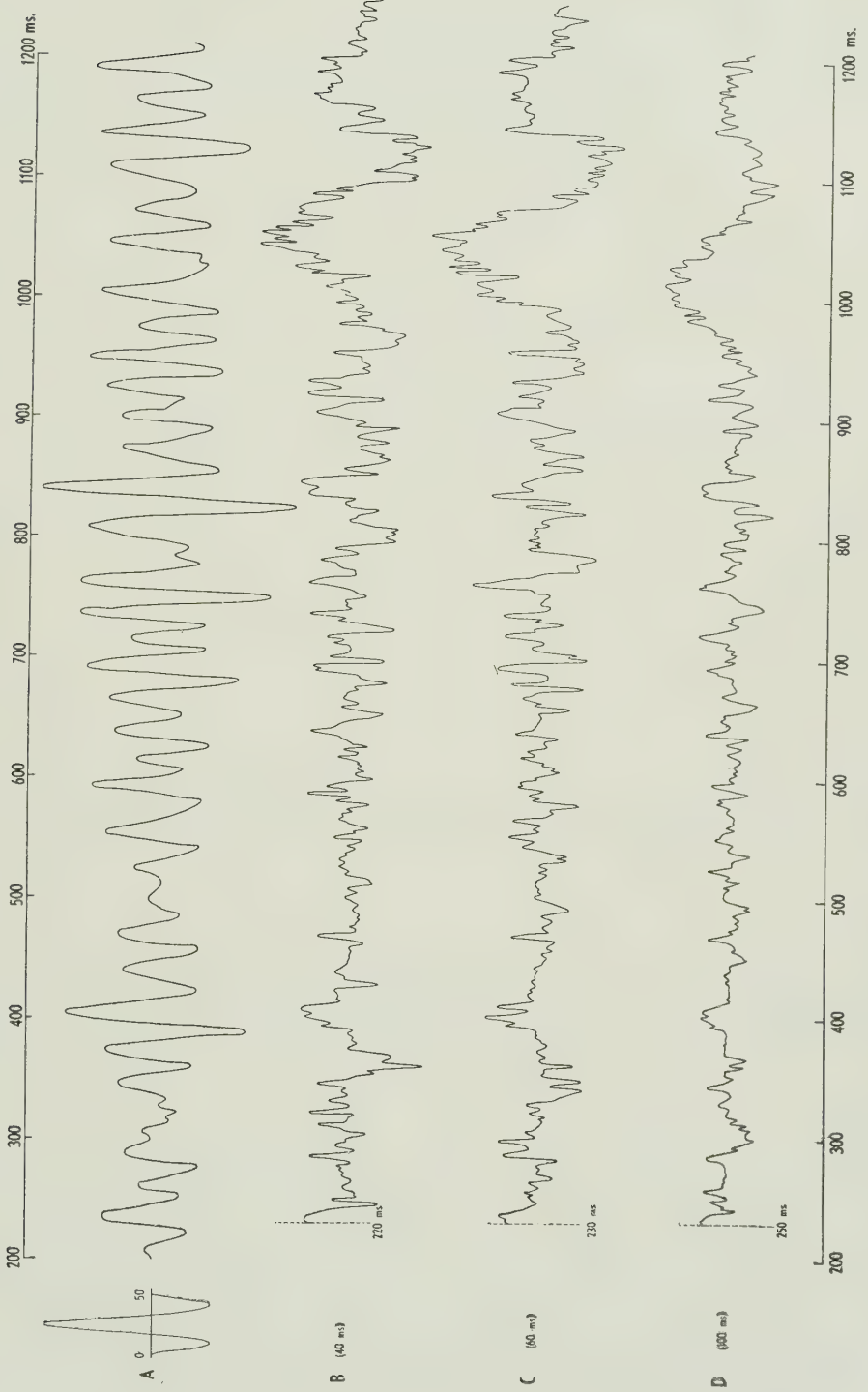
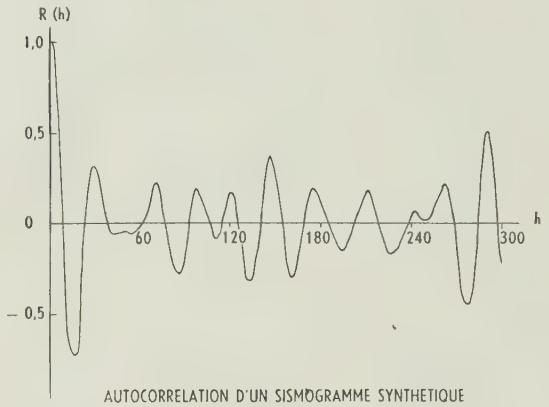
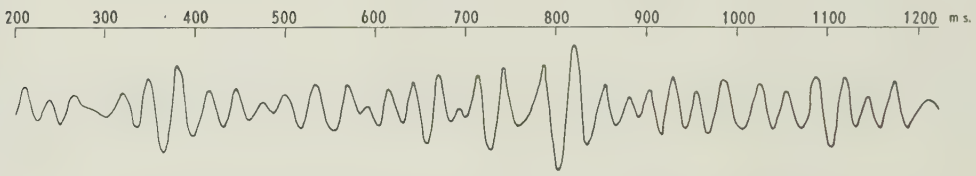


FIG. 23. — Corrélation du sismogramme synthétique avec les coefficients de réflexion moyennés.

A. Sismogramme synthétique.
B. Coefficients de réflexion moyennés sur un intervalle de 40 ms.
C. Coefficients de réflexion moyennés sur un intervalle de 60 ms.
D. Coefficients de réflexion moyennés sur un intervalle de 100 ms.



AUTOCORRELATION D'UN SISMOGRAMME SYNTHETIQUE
(Signal X)



CORRELATION NON NORMEE D'UN SISMOGRAMME SYNTHETIQUE AVEC SON WAVELET GENERATEUR

FIG. 24.

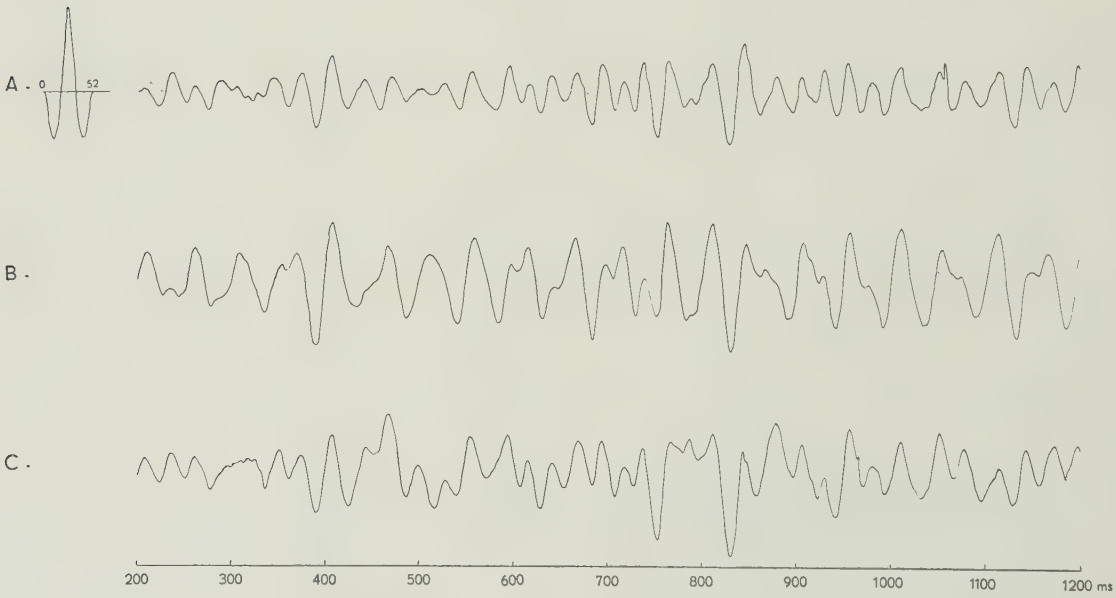


FIG. 25.

A. Sismogramme synthétique.
B. A brouillé par une sinusoïde $f = 20$ c/s amplitude voisine de celle de A.
C. A brouillé par du bruit.

dant à trois impulsions de Ricker allant de 40 à 60 ms. On peut observer dans ce cas le peu de différence entre les trois courbes. Il n'y a par contre presque aucune corrélation à vue entre la courbe des $k\,(t)$ et les courbes $y\,(t)$. Pour faire apparaître une corrélation, il faut filtrer les $k\,(t)$ par une moyenne mobile. La figure 23 montre une mise en phase de la moyenne mobile $k\,(t)$ et de $y\,(t)$ un peu plus nette, bien qu'on observe de très nombreuses différences dues au filtrage des $k\,(t)$, signaux à haute fréquence, par $s\,(t)$ de période 40 à 60 ms.

2. Autocorrélation et corrélation mobile.

Sur la figure 24, nous avons tracé l'autocorrélation du sismogramme synthétique obtenu avec le signal X dans la figure 22. On peut remarquer que cette autocorrélation, calculée sur un secteur de 300 ms, est très loin de s'annuler pour $h > 50$ ms (longueur de $s\,(t)$). Dans la première partie, $\rho\,(h)$ est pratiquement égal à l'autocorrélation théorique d'un Ricker ; mais, sur d'autres exemples, on relève des écarts très importants. Ce fait montre la difficulté qu'il y a à identifier $s\,(t)$ à partir de $\rho\,(h)$.

Sur la figure 24 figure la corrélation mobile de $y\,(t)$ avec l'impulsion d'origine $s\,(t)$. On constate que ce filtrage ne modifie pas trop $y\,(t)$. La corrélation mobile $r\,(t) = \int_0^{\mathfrak{T}} s\,(\tau)\, y\,(t+\tau)\, d\tau$ est seulement déphasée par rapport à $y\,(t)$ de $\frac{\mathfrak{T}}{2}$ dans le cas particulier d'une impulsion de Ricker étudié ici.

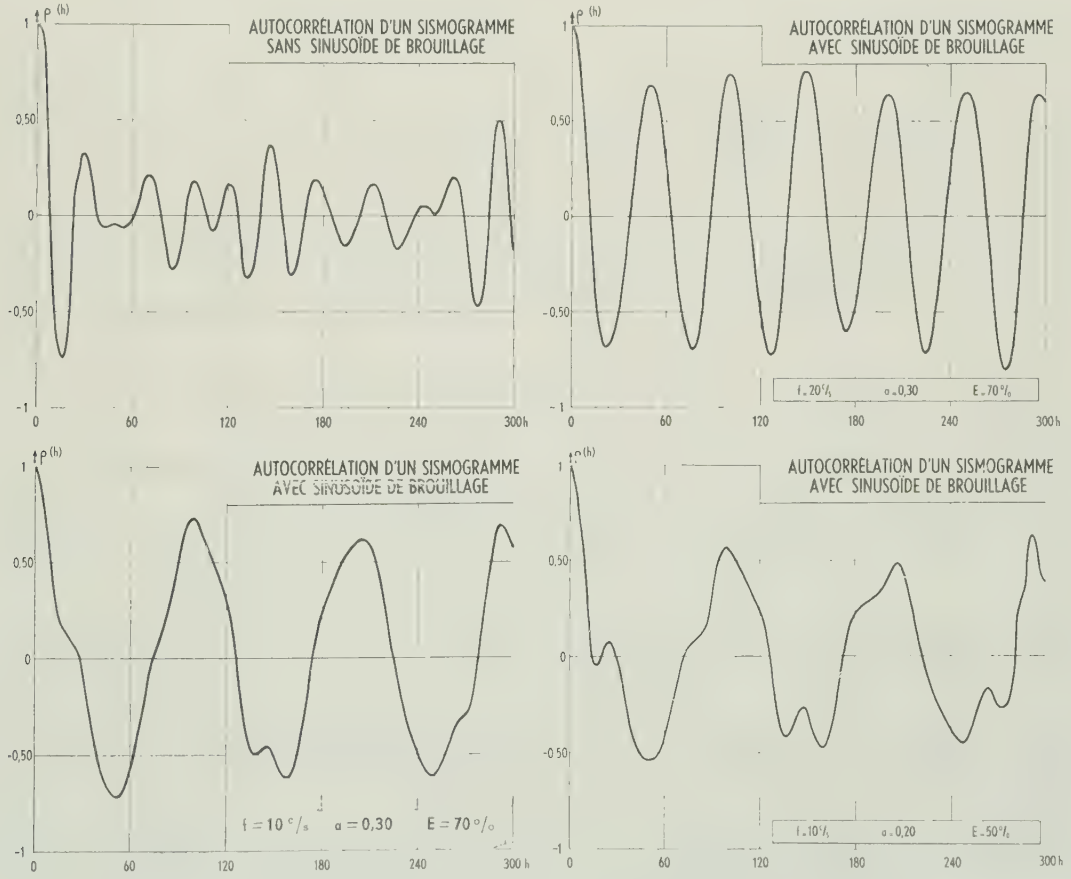


FIG. 26. — Autocorrélations de sismogrammes synthétiques.

3. Élimination du bruit.

On a ajouté au sismogramme synthétique précédent divers brouillages, constitués par des sinusoïdes d'amplitudes et de périodes variables. La figure 25 montre comment un sismogramme synthétique sans multiple est transformé par l'adjonction d'une sinusoïde, ou d'un bruit prélevé sur un sismogramme réel très brouillé. D'une part, le sismogramme initial est très modifié et d'autre part la présence d'une sinusoïde n'est nullement évidente. En revanche, le calcul de l'autocorrélation met bien en évidence cette sinusoïde, comme on le constate sur la figure 26 où nous donnons l'autocorrélation du sismogramme non brouillé et les autocorrélations de ce même sismogramme après brouillage par des sinusoïdes d'amplitudes et de périodes différentes représentant de 50 à 70 % de l'énergie totale du sismogramme brouillé. Les autocorrélations ne sont pas tout à fait de la forme $\rho(h) = E \cos wh$ pour $h > 50$ ms, du fait des résidus $\eta(h)$ qui existent déjà dans le sismogramme brut.

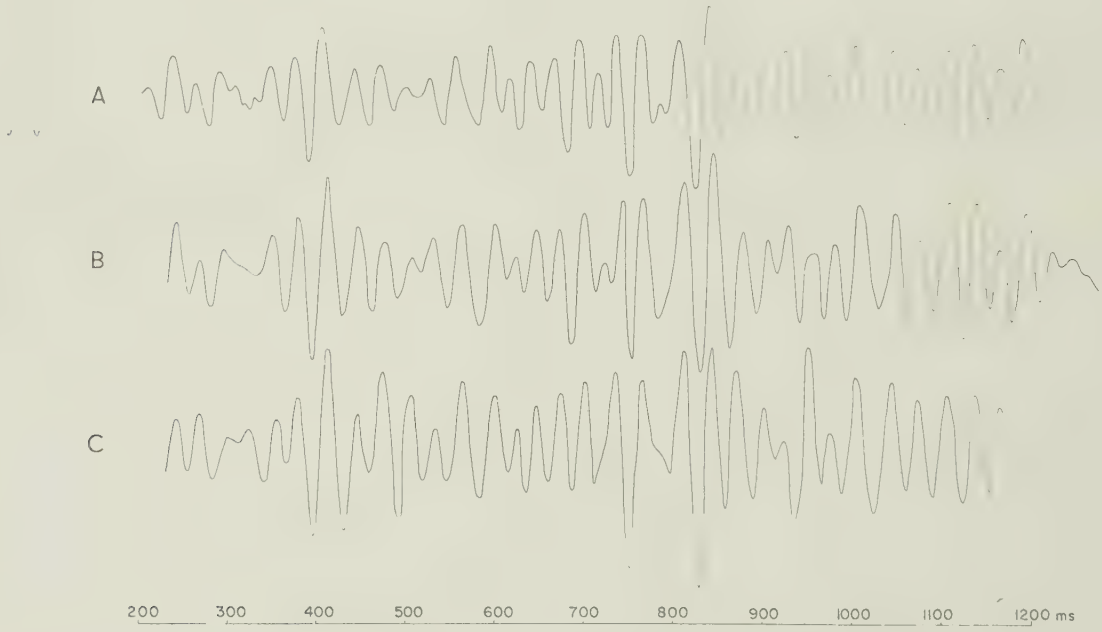


FIG. 27. — La corrélation mobile utilisée comme filtre.

A. Sismogramme synthétique et son implusion X.
B. Corrélation mobile de A brouillé par une sinusoïde avec l'impulsion X.
C. Corrélation mobile de A brouillé par du bruit avec l'impulsion X.

Pour éliminer ces brouillages, on dispose de plusieurs moyens, outre les mixages entre traces, lorsque les bruits ne sont pas en phase. Nous avons employé ici la corrélation mobile du film brouillé par l'impulsion d'origine $s(t)$. On peut voir sur la figure 27 comment la corrélation mobile élimine les bruits signalés sur la figure 25. Le sismogramme brut est restitué presque parfaitement. Remarquons cependant que la détermination de $s(t)$ est laissée de côté et qu'elle semble impossible par l'emploi de l'autocorrélation.

En revanche, l'extraction des sinusoïdes est possible. En effet, l'examen de l'autocorrélation montre les maximums successifs de la sinusoïde $E_0 \cos \omega h$. La valeur moyenne de ces maximums traduit le pourcentage de l'énergie de la sinusoïde dans le signal brouillé. Pour calculer la fréquence et l'am-

plitude de la sinusoïde, il faut calculer l'autocorrélation sur une durée h suffisante pour disposer de trois ou quatre périodes de sinusoïde. On peut estimer ω à vue par une valeur w. Puis, on peut calculer la corrélation de $\rho(h)$ et cos wh entre $\mathcal{H}' < h < \mathcal{H}' + \frac{2k\pi}{w}$ où $\mathcal{H}'$ est supposé supérieur à la longueur $\mathfrak{C}$ des impulsions.

$$R(w) = \int_{\mathcal{H}}^{\mathcal{H}' + \frac{2k\pi}{w}} \rho(h) \cos wh\, dh$$

$$E[R(w)] = E_0 \int_{\mathcal{H}}^{\mathcal{H}' + \frac{2k\pi}{w}} \cos \omega h \cos wh\, dh$$

$R(w)$ est maximal pour $\omega = w$. L'erreur relative commise a été de $\frac{2}{100}$ dans le calcul exposé ici. Ensuite, on estime l'amplitude de la sinusoïde avec :

$$E_0 \int_0^T y^2(t)\, dt = \int_0^T a^2 \sin^2[\omega(t+\varphi)]\, dt = \frac{Ta^2}{2} \text{ si } T \to \infty$$

$$a^2 = \frac{2E_0}{T} \int_0^T y^2(t)\, dt$$

E_0 est la part d'énergie de la sinusoïde dans le film brouillé, et on lit E_0, directement sur l'autocorrélation normée, comme étant la valeur moyenne des minimums et maximums de $\rho(h)$ pour $h = \frac{k\pi}{\omega}$.

Dans notre essai, on a estimé a avec une erreur relative de 5/100.

Il faut ensuite chercher la mise en phase de la sinusoïde $a \sin \omega(t+\varphi)$ avec $y(t)$. Cette opération est faite approximativement à vue et on cherche ensuite le maximum de :

$$r(\varphi) = \int_0^T y(t)\, a \sin \omega(t+\varphi)\, dt$$

La valeur φ_0 maximisant $r(\varphi)$ donne le déphasage optimal. En retranchant $a \sin \omega(t+\varphi_0)$ de $y(t)$, on doit retrouver le sismogramme brut d'origine. L'application de cette méthode a donné d'excellents résultats, mais la présence d'une sinusoïde pure dans un sismogramme étant peu probable, la méthode précédente reste très théorique en sismique. Dans le cas de sismogrammes réels, nous l'avons appliquée, mais les résidus après extraction semblent surtout contenir des bruits et des restes de la fonction périodique qui n'est pas une sinusoïde pure.

Enfin, du fait de la déformation lente de la période, l'extraction n'est satisfaisante que sur un secteur très court.

4. Fabrication de coefficients de réflexion synthétiques.

Pour construire des modèles de sismogramme synthétique, on peut construire des coefficients de réflexion synthétiques. On détermine les coefficients $k(t)$ toutes les 1 ou 2 ms. Pour cela, on part d'une vitesse V (o) donnée correspondant à un palier, puis on calcule les vitesses successives de ce palier en ajoutant à V (o) une variation de vitesse aléatoire de moyenne nulle et obéissant à une loi

normale d'écart type σ compris entre 10 et 50 m/s. Puis, à des instants i obéissant à une loi de Poisson, on procède à un changement de couche géologique en tirant la vitesse V (i) d'un nouveau palier. Cette vitesse obéit à une loi de Gauss dont la moyenne est égale à V (0) ou est une fonction croissante de i et dont l'écart type est de l'ordre de plusieurs centaines de mètres par seconde. Ayant ainsi obtenu une courbe des vitesses synthétiques, on calcule :

$$k(i) = \frac{V(i) - V(i+1)}{V(i) + V(i+1)}$$

IX. ÉLIMINATION DES BRUITS A L'ENREGISTREMENT

Le signal enregistré en surface étant constitué de réflexions profondes arrivant en phase, de bruits inorganisés locaux et d'ondes entretenues périodiques se propageant dans les couches superficielles du sol, on élimine fréquemment les bruits en groupant les mesures de plusieurs sismographes voisins et alignés sur une seule trace. Par ce mixage, les réflexions sont renforcées, les bruits inorganisés amoindris et les bruits organisés périodiques sont presque supprimés dans la mesure où les sismographes d'une trace sont dispersés sur une zone égale à la longueur d'onde du bruit sur le sol ou à un multiple de cette longueur.

Pour étudier les résultats de ces méthodes, l'*Institut Français du Pétrole* a entrepris en 1959 une série de tirs expérimentaux dans deux régions. Lors de chaque tir, on enregistrait sur huit traces les ébranlements du sol au moyen de groupes de neuf sismographes en ligne avec le point de tir et écartés de 5 m par trace, auxquels on ne touchait pas et qui formaient un base de référence. Sur huit autres traces, on enregistrait en même temps les résultats de dispositifs expérimentaux où des lignes de six sismographes étaient regroupées en nombre variable sur une trace. Dans les dispositifs expérimentaux, le nombre de lignes parallèles à la ligne de tir, leur écartement et la distance entre les sismographes d'une ligne étaient variables. Pour comparer les résultats des dispositifs expérimentaux et de référence, on étudie dans un secteur allant de 2 à 3 s, dans lequel les réflexions profondes deviennent très faibles en principe, le rapport des énergies enregistrées. Pour cela on mesure les aires S comprises entre la courbe d'une trace et son amplitude moyenne (fig. 28).

Le rapport des aires des traces de référence et d'expérience indique dans quelle proportion le bruit est éliminé par un dispositif expérimental (23).

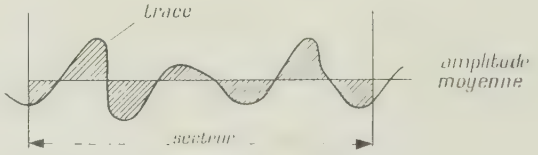


FIG. 28.
Aire hâchurée = S.

1. Étude des traces de référence.

La comparaison des S_{jR_i} (aire d'une trace de référence i lors d'un tir j) et des S_{jE_i} (aire d'une trace expérimentale i lors d'un tir j) nécessite un examen préalable de l'homogénéité de ces populations.

a. Répartition de S_{jR_i}.

Des réglages soignés des gains effectués régulièrement devaient permettre une bonne homogénéité des S_{jRi}. Une étude de la répartition des S_{jR_i} a montré que les S_{jR_i} d'une trace i donnée suivaient

une loi normale, mais que la moyenne $\bar{S}_{R_i}$ et l'écart type σ_{R_i} différaient selon la trace i. Le tableau II fournit les résultats des ajustements pour les deux régions de Saint-Étienne et Gourbera. Des tests χ^2 montrent le degré de validité des ajustements dans le cas de Saint-Étienne. L'existence des nombreuses causes aléatoires lors de la propagation des ondes dans le sol et les différences entre tirs expliquent assez bien la loi de répartition normale. En revanche, les différences entre les moyennes par trace $\bar{S}_{R_i}$, différences très supérieures à $\frac{2\sigma_R}{\sqrt{n}}$ (n : nombre de tirs dans une région) ne s'expliquent pas. On ne peut pas en trouver la raison dans les couplages des sismographes avec le sol [1] car le couplage ne devrait pas donner les mêmes inégalités relatives entre les $\bar{S}_{R_i}$ pour les deux régions, comme on le constate sur la figure 29. On ne peut pas non plus expliquer les différences entre $\bar{S}_{R_i}$ par la distance des sismographes d'une trace au lieu de l'explosion, car les S_{E_i} varient peu autour de leur moyenne générale $\bar{S}_E$ quoique les sismographes des traces expérimentales et de référence numéro i soient à une même distance moyenne de l'explosion.

Nous avons donc conclu que les S_{jR_i} n'étaient pas homogènes. Les gains des diverses traces sont inégaux, ce qui explique les écarts des moyennes et des écarts types. Pour corriger ces écarts de gain, on a corrigé les S_{jR_i} des traces i dont les moyennes $\bar{S}_{Ri}$ s'écartaient de plus de deux écarts types $\frac{\sigma_R}{\sqrt{n}}$ de la moyenne générale $\bar{S}_R$ en multipliant S_{jRi} par $\frac{\bar{S}_R}{\bar{S}_{R_i}}$. On a ainsi modifié les traces 3, 4, 5 et 7 de Saint-Étienne et 2, 3, 5, 7 de Gourbera. Après cette opération, on obtient des S'_{jR_i} recentrés.

On peut constater sur le tableau II que les écarts types des S_{R_i} d'une trace diffèrent significativement. Dans le tableau nous avons marqué d'un astérisque les valeurs sortant de l'intervalle de confiance à 95 %, c'est-à-dire, telles que :

$$\bar{S} - \frac{2\sigma}{\sqrt{n}} < S_i < S + \frac{2\sigma}{\sqrt{n}}$$

$$\sigma - \frac{2\sigma}{\sqrt{2n}} < \sigma_i < \sigma + \frac{2\sigma}{\sqrt{2n}}$$

$\left(\frac{\sigma}{\sqrt{2n}}\right.$ étant l'écart type de $\left.\sigma_i\right)$.

Nous n'avons rien fait pour remédier au manque d'homogénéité des σ. On a constaté que le recentrage des S_{jR_i} réduisait la variance due aux tirs de moitié dans le cas de Saint-Étienne ($\sigma_R^2 = 33^2$ et $\sigma'^2_R = 22{,}9^2$) et de 30 % dans le cas de Gourbera. Par contre, les valeurs $\frac{S_{jE}}{S'_{jR}}$ ne sont presque pas modifiées par les corrections.

En ce qui concerne la répartition globale des S'_{jR_i} elle ne s'ajuste pas bien à une loi normale, du fait de la diversité des σ_{R_i} des huit traces. Il en est de même pour l'ensemble des S_{jE_i}, les différences entre les S_{jE_i} n'étant pas complètement aléatoires puisqu'on espère qu'elles proviennent en grande part des dispositifs expérimentaux et de leur efficacité dans le filtrage du bruit.

b. Analyse de la variance des S'_{jR_i}.

Nous avons effectué une analyse de la variance sur les valeurs S'_{jR_i} afin de mesurer les causes de dispersion dues aux tirs et aux traces. On obtient pour Saint-Étienne :

---

[1] A moins que les zones de plantage des sismos de référence de la trace 7 ne soient dans une zone d'excellent couplage.

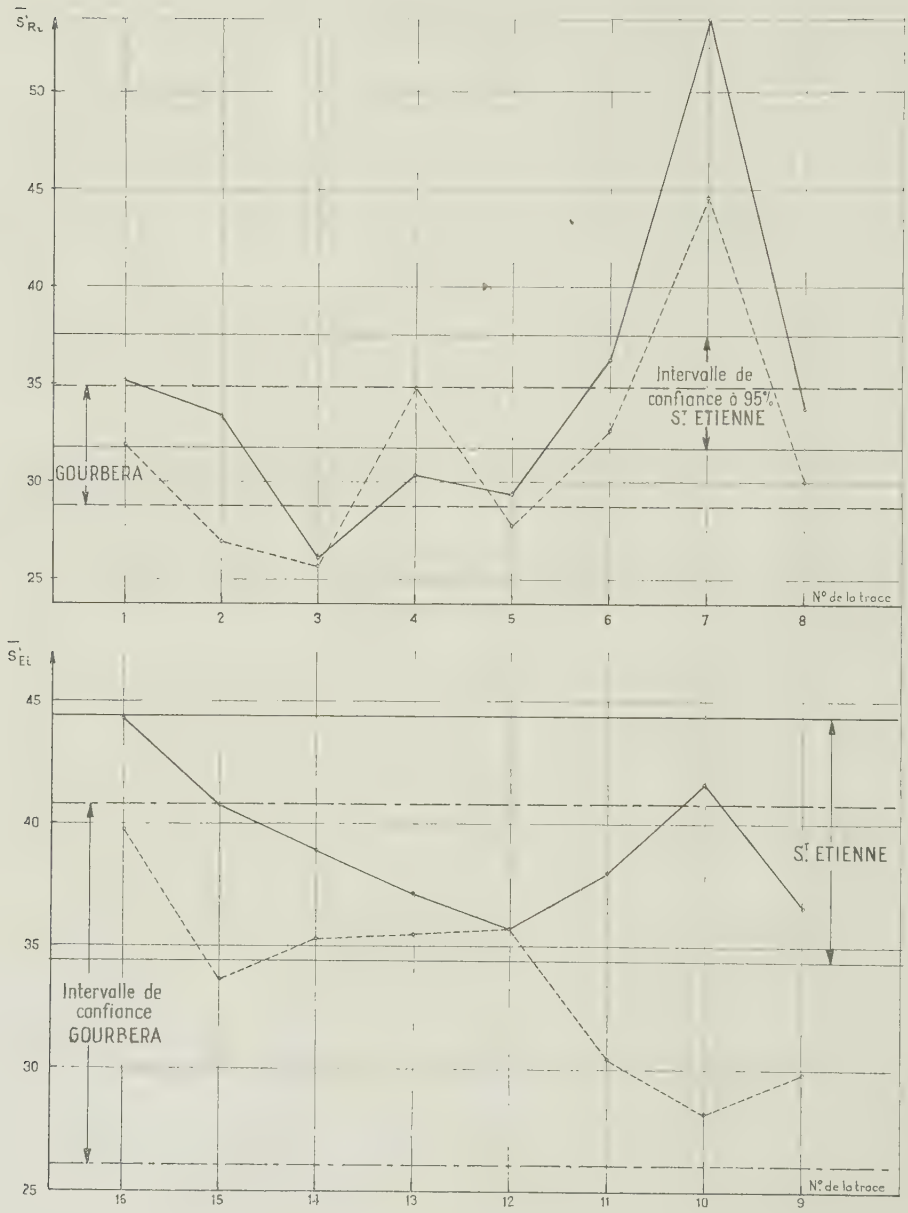


FIG. 29. — Courbes des $\overline{S}_{R_i}$ et des $\overline{S}_{E_i}$.

SAINT-ÉTIENNE — –
COURBERA - - - - -

Tableau II.

Analyse des S_{jRi} et S_{jEi}.

		MOY.	TRACE 1	TRACE 2	TRACE 3	TRACE 4	TRACE 5	TRACE 6	TRACE 7	TRACE 8
$\overline{S}_{Ri}$	Saint-Étienne	31,8	35,2	33,3	26,1*	30,4*	29,3*	36,2	53,6*	33,7
σ_{Ri}	—	12,1	11,6	7,3	5,8	5,8	6,5	6,4	16,5	7,4
$\overline{S}'_{Ri}$	—	34,6	35,2	33,3	34,5	34,6	34,6	36,2	34,8	33,7
σ'_{Ri}	—	8,4	11,6*	7,3	7,8	6,4*	7,5	6,4*	10,7*	7,4
χ^2	— %	2,5	12	35	66	20	74	45	80	42
$\overline{S}_{Ei}$	—	39,4	44,3	40,7	38,9	37,1	35,7	38,0	41,6	36,6
σ_{Ei}	—	21,9	24,6	20,1	24,4	20,7	16,9*	20,4	24,1	16,9*
$\overline{S}_{Ri}$	Gourbera	31,8	32,0	26,9*	25,6*	34,7	27,8*	32,7	44,7*	30,0
σ_{Ri}	—	11,6	11,5	8,1	8,0	14,8	8,8	8,2	11,3	10,6
$\overline{S}'_{Ri}$	—	31,7	32,0	32,2	32,3	34,7	32,0	32,7	32,3	30,0
σ'_{Ri}	—	10,4	11,5	8,7	8,9	14,8*	10,2	8,2*	8,0*	10,6
$\overline{S}_{Ei}$	—	33,5	39,7	33,6	35,3	35,4	35,6	30,4	28,2	29,8
σ_{Ei}	—	28,6	34,1*	24,6	29,5	29,3	33,7	26,0	22,3*	25,3

(*) L'astérisque indique que la valeur correspondante sort de l'intervalle de confiance à 95 % d'une variable normale.

Intervalles de confiance a 95 %.

	SAINT-ÉTIENNE	GOURBERA
S'_R	32,6 à 36,6	29,0 à 34,5
S_E	34,4 à 44,5	26,1 à 40,9
σ'_R	6,9 à 9,8	8,4 à 12,4
σ_E	18,3 à 25,4	23,3 à 34,0

s_1^2 variance inter explosions : $$\frac{n \sum_j (\bar{S}'_{jR} - S'_R)^2}{8 - 1} = 2\,484$$

s_2^2 variance inter trace : $$\frac{8 \sum_i (S'_{R_i} - S'_R)^2}{n - 1} = 687$$

s_3^2 variance résiduelle : $$\frac{\sum_{ij} (S'_{jR_i} - S'_{R_i} - S'_{jR} + S'_R)^2}{(n - 1)\ (8 - 1)} = 244$$

s_4^2 variance totale : $$\frac{\sum_{ij} (S'_{jR_i} - \bar{S}'_R)^2}{8n - 1} = 526$$

s_5^2 variance intra trace : $$\frac{\sum_{ij} (S'_{jR_i} - \bar{S}'_{R_i})^2}{8n - n} = 569$$

s_6^2 variance intra explosions : $$\frac{\sum_{ij} (S'_{jR_i} - \bar{S}'_{R_i})^2}{8n - 8} = 250$$

L'application du test de Snédécor indique une variation significative des énergies due à l'explosion $\left(f_{10} = \frac{s_1^2}{s_3^2} = 10 \text{ avec } 1\ \% \text{ de chances pour que } f > 1\right)$. L'effet de la trace existe encore après correction $\left(f_{23} = \frac{s_2^2}{s_3^2} = 2{,}80 \text{ et } 1\ \% \text{ de chances que } f_{23} > 2{,}8\right)$, mais, avant correction des S_{jR_i}, cet effet était prépondérant puisque f_{23} égalait 55 au lieu de 2,8.

Le rapport $f_{61} = \frac{s_1^2}{s_6^2} = 9{,}9$ (f_{61} ayant seulement 1 % de chances de dépasser 1) ainsi que le rapport $f_{52} = \frac{s_2^2}{s_5^2} = 1{,}2$ (f_{52} a 5 % de chance de dépasser 2,1) montrent qu'après correction sur les S_{jR_i} la seule cause significative de variation des S'_{jR_i} est l'explosion.

c. Corrélation entre traces d'un film.

En normant S'_{jR_i} par S_{jR} pour éliminer l'influence de l'explosion dans un tir, on peut étudier les couples $\left(\frac{S'_{jR_i}}{\bar{S}'_{jR}}, \frac{S'_{jR_{i+l}}}{\bar{S}'_{jR}}\right)$ pour déterminer s'il existe une corrélation des énergies reçues entre traces voisines. On a construit les nuages des couples relatifs à vingt tirs, avec $l = 1, 3, 5$ et 7, sans déceler de corrélation. La plus forte corrélation correspondait aux couples de traces consécutives ($l = 1$). Le coefficient de corrélation linéaire était alors $r = -0{,}12$, soit $r^2 = 1{,}4$ %. Ce coefficient n'est pas significatif. Le rapport f de Suédécor est égal à 1,9 alors qu'il a 5 % de chances d'être supérieur à 3,8 du seul fait du hasard.

d. Variation de l'énergie.

Contrairement aux suppositions de stationnarité nécessaires pour mener à bien des calculs théoriques, nous avons constaté en découpant les S_{jR_i} en deux secteurs (l'un de 2 à 2,5 s, l'autre de 2,5 à 3 s) que l'énergie reçue diminue presque de moitié d'un secteur à l'autre.

2. Résultats de l'expérimentation.

Nous examinons ici les rapports $\frac{S_{jE_i}}{S'_{jR_i}}$ ou $\frac{\bar{S}_{jE}}{\bar{S}'_{jR}}$ afin d'étudier le filtrage du bruit résultant des dispositifs expérimentés. En normant par S'_{jR}, on doit en principe éliminer les variations importantes d'énergie liées au tir. Les tableaux III et IV retracent l'ensemble des valeurs $\frac{\bar{S}_{jE}}{\bar{S}'_{jR}}$, en fonction des trois paramètres étudiés (distance entre sismographes ; nombre de lignes ou nombre de sismographes ; distance entre lignes).

Le nombre de sismographes par trace est égal au produit du nombre de lignes par 6 (nombre de sismos par ligne).

L'exploitation des résultats des tableaux III et IV est délicate car les paramètres varient simultanément.

a. Largeur du dispositif.

Elle est égale à $(k-1)\,l$ où k est le nombre de lignes et l la distance entre lignes. La figure 30 montre la diminution moyenne du bruit avec la largeur croissante du dispositif. L'élargissement du dispositif ne fait guère gagner plus de 25 % pour un même nombre de sismographes utilisés dans l'expérience.

b. Longueur du dispositif.

Les figures 31 et 32 montrent la diminution du bruit avec l'accroissement de la longueur du dispositif pour des nombres de sismographes donnés. Il est regrettable que l'on n'ait pas utilisé de dispositif plus long. Le bruit est divisé par 4 lorsque la longueur croît de 1 à 75 m, dans le cas de Saint-Étienne. A Gourbera, il semble que le bruit passe par un minimum pour une longueur de 25 à 50 m.

c. Nombre de sismographes, ou nombre de lignes.

Dans un article (23), Richard et Rimbaut indiquaient que le bruit décroissait comme $\frac{K}{\sqrt{n}}$ fois le nombre n de sismographes par trace. Ce raisonnement valable pour des bruits indépendants des sismographes risque de ne plus convenir pour l'élimination d'un bruit périodique. Dans ce dernier cas, un dispositif approprié de quelques sismographes a autant d'effet qu'un grand nombre de sismographes dispersés.

En moyenne, les expériences dont nous disposons semblent indiquer une certaine décroissance du bruit en fonction de l'accroissement des sismographes, mais on ne gagne que de 30 à 60 % en passant de six à quatre-vingt-seize sismographes par trace au lieu de gagner 75 % selon la formule théorique. Dans les régions étudiées et pour les dispositifs de « couverture latérale » utilisés, l'influence du nombre de sismographes semble beaucoup moins importante que celle de la longueur du dispositif. Dans la mesure où un bruit organisé existe, il apparaît du reste évident que c'est la disposition et non le nombre de sismographes qui élimine le bruit.

Tableau III.

Tableau des valeurs de $\frac{\overline{S_E}}{\overline{S_R}}$ par film pour Saint-Étienne.

ÉCARTEMENT LIGNES (m)			0,20	1					5			10			15			20			30		
ÉCARTEMENT SISMOS (m)			0,20	1	3	5	10	15	5	10	15	5	10	15	5	10	15	5	10	15	5	10	15
NOMBRE DE LIGNES	1	NOMBRE DE SISMOS 6	2,63	3,15	2,38	1,70	1,45	0,42															
	2	12		2,99	1,70	1,42	1,21	0,30	1,12	0,82	0,88	1,07	0,77	0,90	1,15	0,88	0,74	0,93	0,82	0,82	0,88	0,79	0,77
	4	24	2,36	2,82	1,59	1,18	1,21	0,77	0,90	1,07	0,74	0,93	0,79	1,01	0,71	0,82	0,77	0,82	0,74	0,71	0,79	0,68	0,77
	8	48	2,69	2,31	1,23	0,99	0,77	0,64	0,71	0,71	0,74	0,71	0,74	0,68	0,60	0,47	0,52						
	12	72	2,22	2,49	1,70	1,20	0,88	0,80															
	16	96		2,22	1,67	1,01	0,85	0,75															

TABLEAU IV.

TABLEAU DES VALEURS DE $\frac{\overline{S_E}}{\overline{S_R}}$ PAR FILM POUR GOURBERA.

NOMBRE DE LIGNES	DISTANCE (m) LIGNES	DISTANCE (m) SISMOS	0,2	1	3	5	10	15	5	10	15	5	10	15	5	10	15	5	10	15	5	10	15
			1																				
1		6	2,83	2,67	1,63	1,71	0,93	1,71															
2		12	2,72	2,26	1,48	1,05	0,88		0,93	1,04	1,21	0,96	1,03	1,12	0,80	0,79	0,76	1,03	0,88	1,11	0,79	0,64	0,90
4		24	2,75	2,07	1,07	0,28	0,93	0,63	0,67	0,75	1,03	0,65	0,70	0,84	0,49	0,57	0,56	0,57	0,59	0,60	0,64	0,54	0,59
8		48	2,03	1,78	0,97	0,61	0,46	0,70	0,39	0,44	0,60	0,43	0,52	0,58	0,41	0,42	0,41						
12		72	1,52	1,39	0,86	0,61	0,99																
16		96	1,40	1,17	0,68	0,45	0,72	1,03															

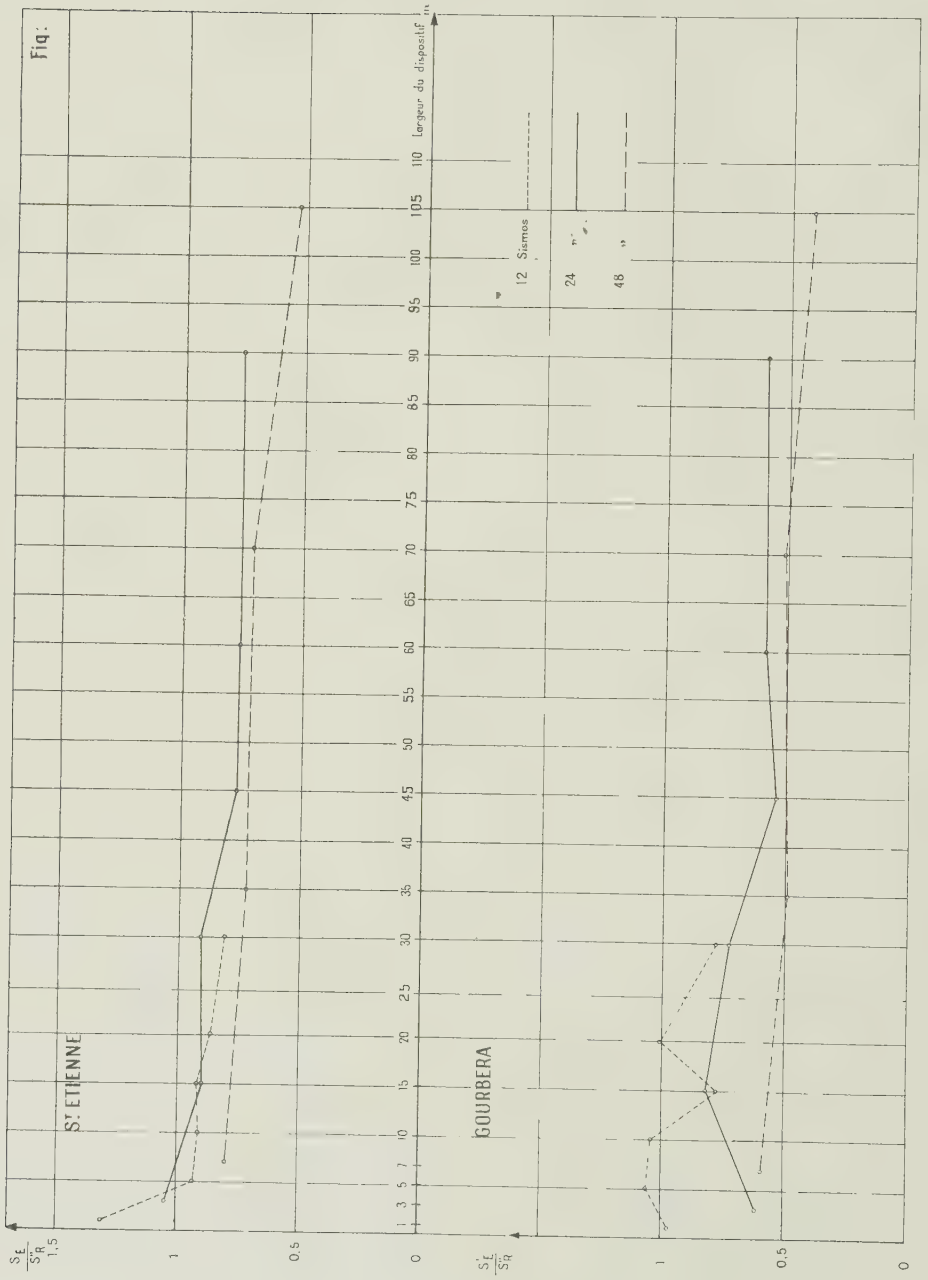


FIG. 30. — Influence de la largeur du dispositif.

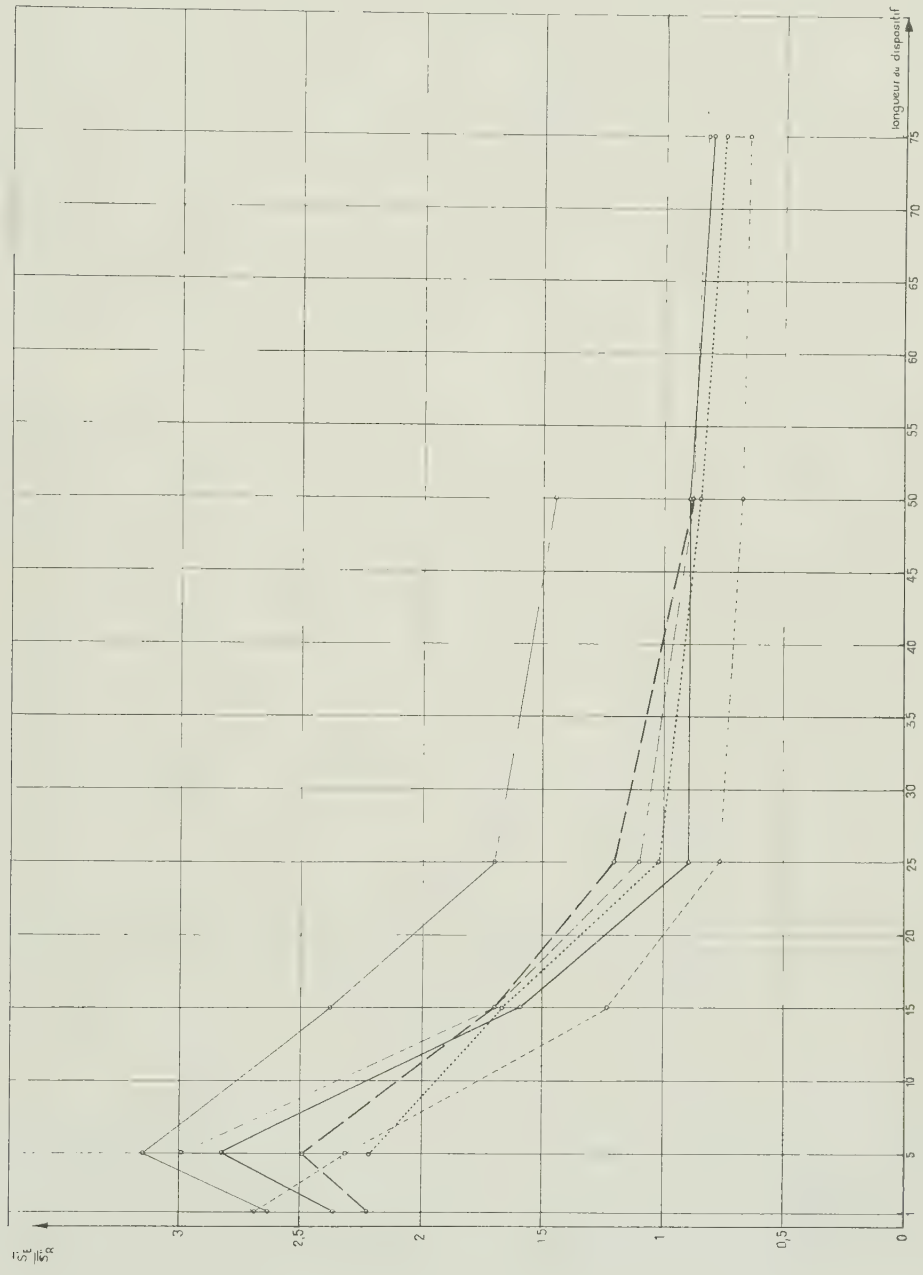


FIG. 31. — Influence de la longueur du dispositif. ST-ÉTIENNE.

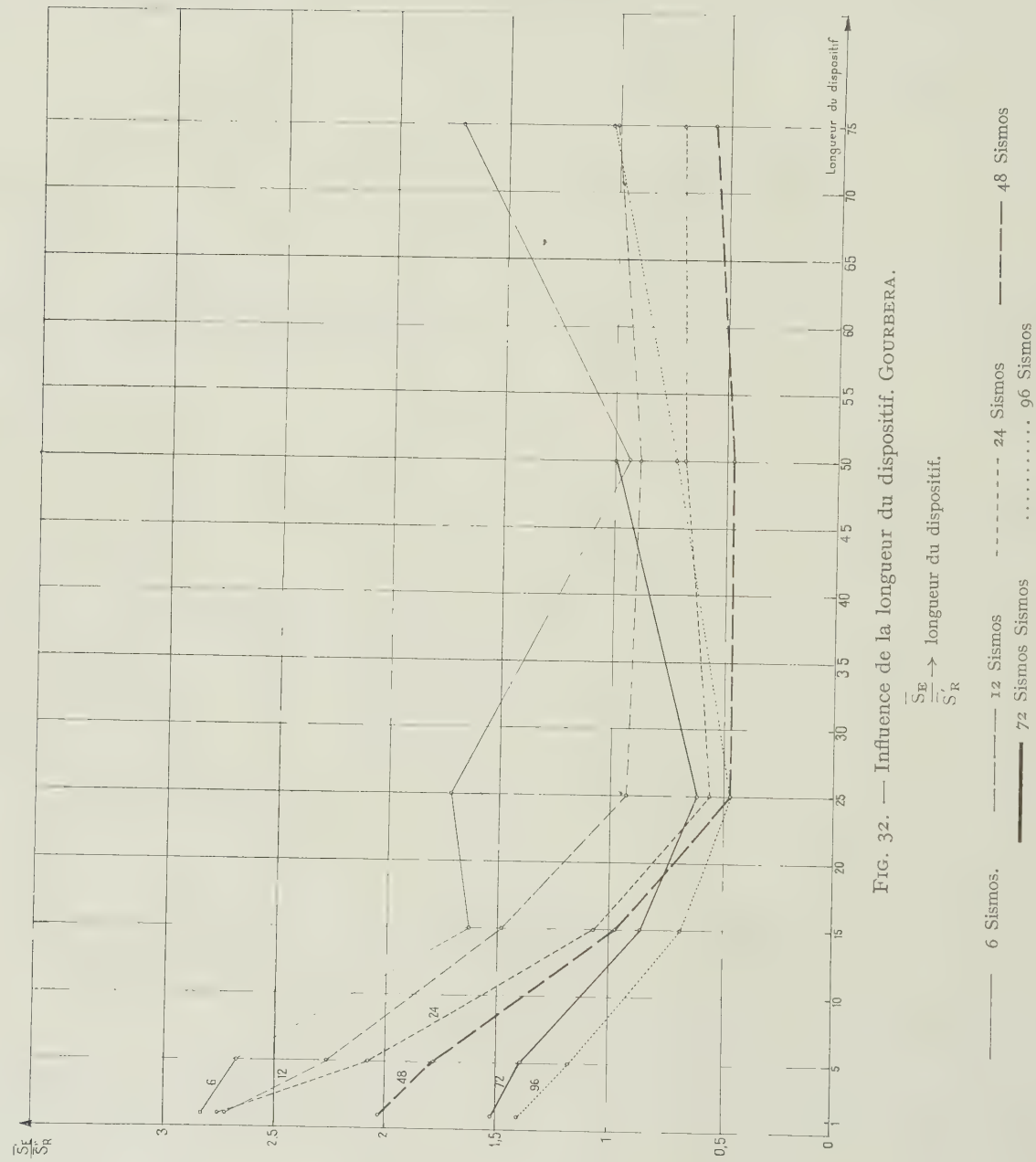


FIG. 32. — Influence de la longueur du dispositif. GOURBERA.

$\frac{\overline{S_E}}{\overline{S'_R}}$ → longueur du dispositif.

——— 6 Sismos. — — — — 12 Sismos ---------- 24 Sismos — — — — 48 Sismos

——— 72 Sismos Sismos ..... 96 Sismos

d. Dispersion des résultats par tir.

Nous avons étudié les rapports $\frac{S_{jE_i}}{S'_{jR_i}}$ tir par tir, et nous avons mesuré « l'étendue », écart entre les rapports extrêmes $\frac{S_{jE_i}}{S'_{jR_i}}$ d'un tir j. On constate que l'étendue diminue lorsque la longueur du dispositif ou le nombre de sismographes croissent. D'autre part, les recentrages apportés aux S_{jR_i} de certaines traces diminuent l'étendue d'environ 10 %.

3. Analyse des amplitudes.

L'analyse des énergies nous a semblé insuffisante pour mesurer l'effet des dispositifs expérimentaux sur le bruit. C'est pourquoi nous avons fait relever les amplitudes de quelques traces échantillonnées sur les divers dispositifs pour étudier les amplitudes, leur dispersion, les autocorrélations et intercorrélations, qui nous échappent dans la méthode précédente.

a. Répartition des amplitudes.

Les amplitudes ont été relevées de 5 en 5 ms sur certaines traces des dispositifs expérimentaux et sur cinq traces de référence d'un même film. Les amplitudes suivent parfaitement des lois normales de moyenne nulle approximativement et dont les écarts types σ figurent sur les tableaux V et VI. On remarque que le rapport de σ à S est pratiquement constant $\left(\frac{\sigma}{S}\right.$ varie de 0,22 à 0,25, compte tenu des unités adoptées dans les mesures respectives). Ce fait semble assez naturel car la surface S, comme l'écart type σ sont directement proportionnels à l'énergie du bruit enregistré.

b. Étude des autocorrélations.

Les tableaux V et VI récapitulent les résultats des analyses des films étudiés. On constate que les périodes principales, estimées directement à vue sur la trace, coïncident en général avec la période ressortant de l'autocorrélation. D'autre part, la périodicité des maximums et des minimums de l'autocorrélation dans de nombreux cas semble indiquer la présence d'un bruit organisé périodique représentant 20 à 40 % et plus de l'énergie portée par la trace. En revanche, il est très décevant de constater que la longueur du dispositif semble sans influence sur l'élimination de ce bruit organisé périodique qu'elle devrait contribuer à faire disparaître. La période moyenne de 60 ms observée correspond sur le terrain à une longueur d'onde de 30 à 60 m avec des vitesses superficielles de 500 à 1 000 m/s. De ce fait, les dispositifs à plus de 50 m devraient éliminer les ondes de surface, ce qui ne semble pas le cas.

c. Intercorrélations.

Nous avons calculé les intercorrélations entre les traces 1-2, 1-4, 1-6, 1-8 sur deux films de référence. L'intercorrélation fait apparaître en général une période de 50 à 60 ms, comme les autocorrélations (tableaux V et VI). Selon le cas, l'intercorrélation normée peut passer par des maximums de 25 à 50 %. Ces intercorrélations semblent indiquer une corrélation aléatoire entre traces. En effet, avec un déphasage nul, les intercorrélations ne passent pas par un maximum. Or, au bout de deux secondes d'enregistrement, les réflexions, s'il y en a, devraient arriver en phase. Il ne semble donc pas qu'on puisse supposer l'existence de réflexions importantes. D'autre part l'existence d'ondes entretenues périodiques en surface devrait se constater par une mise en phase des intercorrélations des traces grâce à des déphasages croissants, contrairement à ce que montrent les intercorrélations.

TABLEAU V.

SAINT-ÉTIENNE

	FILM N°	TRACE N°	NOMBRE SISMOS	ÉCART. SISMOS (m)	LONG. DISPO-SITIF (m)	σ AMPL.	S_E	S_R	σ/S	PÉRIODES τ ms film	PÉRIODES τ ms auto-corrélation	A* (%)
AUTOCORRÉLATION	168	14	6	1	5	28	111		0,25	80	83	40
	163	16	12	3	15	45	158		0,28	66	70	40
	149	13	24	5	25	23	99		0,23	60	100	60
	196	16	72	10	50	51	189		0,27	63	61	50
	150	1	9	5	40	26		117	0,22	59	65	30
INTERCORRÉLATION	150	1–2	9	5	40	28		107	0,25	59		
	150	1–4	9	5	40	21		91	0,23	52	59	
	150	1–6	9	5	40	28		135	0,22	77		
	150	1–8	9	5	40	31,5		126	0,25	80		

* A : valeur moyenne des extremums en valeur absolue.

TABLEAU VI.

GOURBERA

	FILM N°	TRACE N°	NOMBRE SISMOS	ÉCART. SISMOS (m)	LONG. DISPOSITIF (m)	σ AMPLIT.	S_E	S_R	σ/S	PÉRIODES τ ms film	PÉRIODES τ ms auto-corrélation	A* (%)
AUTOCORRÉLATION	158	16	24	0,2	1	40,6	180		0,23	32 à 45	30 à 40	
	155*a*	15	12	1	5	34,6	152		0,23	35	60	
	174	16	72	3	15	31,7	130		0,24	63	63	40
	176	14	72	5	25	30,4	125		0,24	60	65	40
	182	16	72	10	50	75,3	325		0,23	57	55	80
	227	16	12	10	50	25,3	118		0,22	35 à 27	60 à 12	30
	183	12	96	15	75	42,4	180		0,24	32	57	20
	197	12	48	15	75	19,7	78		0,25	30	59	40
	172	1	9	5	40	25		106	0,25		30	
INTERCORRÉLATION	172	1–2	9	5	40	20,9		86	0,24	50		
	172	1–4	9	5	40	35,2		140	0,25		30	
	172	1–6	9	5	40	22		93	0,24			
	172	1–8	9	5	40	23		99	0,23			

* A : valeur moyenne des extremums en valeur absolue.

X. DÉTECTION AUTOMATIQUE DES RÉFLEXIONS

Actuellement, la lecture des sismogrammes est faite à vue, c'est-à-dire que des « pointeurs » observent les sismogrammes développés sur papier et recherchent des alignements caractérisant une réflexion. On peut essayer d'automatiser ce travail à l'aide d'une machine traitant directement des enregistrements magnétiques sur lesquels le contrôle automatique de gain n'est plus nécessaire du fait de la grande sensibilité de ce genre d'enregistrement.

1. Caractéristiques d'une réflexion.

Une réflexion est caractérisée par :

— *l'identité de la forme* sur un ou deux cycles pour l'ensemble des traces (au moins 75 % d'entre elles) ;

— *l'amplitude* et les écarts entre maximum et minimum de la réflexion doivent être d'importance analogue sur toutes les traces ;

— *l'enveloppe*, c'est-à-dire les courbes joignant les maximums d'une réflexion sur les traces d'une part et les minimums d'autre part doit être continue. Sa forme doit recouper l'enveloppe observée sur les réflexions voisines sauf si le pendage de la couche réfléchissante considérée n'est pas parallèle aux autres pendages.

2. Probabilité d'apparition d'une réflexion.

Une réflexion peut être dissimulée par le bruit.

Soit B (x) la probabilité que le bruit soit supérieur à x. Supposons qu'une réflexion ait son maximum en t sur toutes les traces. Pour que cette réflexion soit admise sur une trace i, il faut que son maximum soit situé entre $t-\tau$ et $t+\tau$ sur la trace i. La quantité τ traduit notre tolérance dans l'acceptation d'une réflexion au vu de l'alignement de son maximum sur les diverses traces (en général les pointeurs admettent environ $\tau = 5$ ms). Soit S l'amplitude moyenne du maximum de la réflexion sur les traces. On peut estimer que cette réflexion sera cachée sur une trace dans deux cas : si le bruit x est inférieur à $-\frac{S}{2}$ entre $t-\tau$ et $t+\tau$ ou si le bruit x est supérieur à S à l'extérieur de $(t-\tau,\ t+\tau)$. Étant donné la continuité du bruit excluant la simultanéité des deux possibilités, on peut estimer que la probabilité d'élimination de la réflexion est :

$$P(S) = B(S) + 1 - B\left(-\frac{S}{2}\right)$$

En admettant que les lois des amplitudes du bruit sont les mêmes sur toutes les traces, la probabilité d'apparition P_A de la réflexion sur j traces parmi les n traces du film sera :

$$P_A = C_n P^{n-j}(S)\,[1 - P(S)]^j$$

La probabilité d'apparition de la réflexion sur j_1 traces au moins sera :

$$P_{A_1} = \sum_{j=j_1}^{n} C_n^j\, P^{n-j}(S)\,[1 - P(S)]^j$$

On peut rapidement calculer P_{A_1} en admettant que B (S) suit une loi de Gauss de moyenne nulle et d'écart type σ, et en utilisant les Tables of the cumulative binomial probability distribution (1).

3. Probabilité d'une fausse réflexion.

Si nous nous plaçons dans une zone où aucune réflexion n'existe, on peut se demander quelle est la probabilité de confondre des bruits avec une réflexion. Nous supposons que les bruits ne sont pas organisés, mais qu'ils ont un spectre voisin de celui des réflexions, ce qui est le cas courant. Comme on peut admettre que la loi de probabilité du bruit est stationnaire, la probabilité que le bruit de la trace i passe par un maximum entre $t - \tau$ et $t + \tau$ est égale à $\frac{2\tau}{\mathfrak{T}}$ où $\mathfrak{T}$ est la période moyenne du bruit.

La probabilité que j_1 maximums aléatoires ou plus coïncident parmi les n traces dans la zone $t - \tau$, $t + \tau$ est :

$$P_F = \sum_{j=j_1}^{n} C_n^j \left(\frac{2\tau}{\mathfrak{T}}\right)^j \left(1 - \frac{2\tau}{\mathfrak{T}}\right)^{n-j}$$

Sur un sismogramme de longueur T où l'on peut observer $\frac{T}{\mathfrak{T}} \simeq m$ maximums, on voit que le nombre moyen des fausses réflexions sera :

$$P = m \sum_{j=j_1}^{n} C_n^j \left(\frac{2\tau}{\mathfrak{T}}\right)^j \left(1 - \frac{2\tau}{\mathfrak{T}}\right)^{n-j}$$

4. Distribution de probabilité des réflexions.

Les deux calculs précédents simplifient beaucoup trop les caractéristiques de la réflexion. Dans un article de *Geophysics* (20), L. Massé propose une méthode plus élaborée. Il considère un film sur lequel il suppose l'existence de une ou plusieurs réflexions avec des probabilités a priori $P_{r_0}(t_r)$. En examinant la première trace, on trouve une probabilité $L_{r_1}(t_r)$ qu'une réflexion se trouve au point t_r. La probabilité de la réflexion en t_r devient après examen de la première trace :

$$P_{r_1}(t_r) = k_1 P_{r_0}(t_r) L_{r_1}(t_r)$$

En continuant de proche en proche, on trouve après examen de m traces

$$P_{r_m}(t_r) = k P_{r_0}(t_r) \prod_{i=1}^{m} L_{r_i}(t_r)$$

Massé propose de prendre :

$$L_{r_i}(t_r) = C\, e^{\frac{2}{N}\int_0^{\mathfrak{T}} s(\tau)\, y_i(t+\tau)\, d\tau}$$

où $s(\tau)$ est l'impulsion émise et $\frac{N}{2}$ la demi-énergie du bruit. Si le nombre de traces m est très grand, Massé dit que : $\lim P_{r_m}(t) \to \delta(t - t_r)$, fonction de Dirac. La fonction $P_{r_m}(t)$ indique des pics aux emplacements des réflexions et des valeurs faibles ailleurs.

(1) *Harvard University Press*, Cambridge.

La méthode ne permet cependant pas de déterminer l'impulsion $s\ (t)$, elle fait un emploi arbitraire de la forme de L_{tr}, et la normalisation $\frac{N}{2}$ de même que le $P_{r_0}\ (t_r)$ sont arbitraires.

D'autre part, la probabilité finale est :

$$P_{r_m}\ (t_r) = C\, e^{\int_0^{\infty} s(\tau) \left[\sum_i^m y\ (t_r + \tau) \right] d\tau}$$

C'est tout simplement une transformée de la corrélation mobile de la moyenne des traces par $s\ (\tau)$, comme l'indique l'auteur. Dans les cas où une réflexion ne revient pas au même instant t_r sur chaque trace (cas des premières réflexions et des couches non horizontales) la méthode élimine la réflexion.

5. Détection automatique des réflexions.

Pour parvenir à une détection automatique des réflexions, on pourrait, à partir des enregistrements magnétiques des films, calculer la moyenne de l'autocorrélation des traces sur chaque secteur et en tirer une impulsion constituée par le centre de l'autocorrélation (fig. 33).

La corrélation mobile $r\ (t)$ de ce signal, généralement assez voisin de l'impulsion réelle $s\ (t)$ qui n'est cependant pas toujours symétrique, permettrait éventuellement d'éliminer certains bruits. La corrélation $r\ (t)$ serait en moyenne en avance de r sur $y\ (t)$ et il serait facile de les remettre en phase.

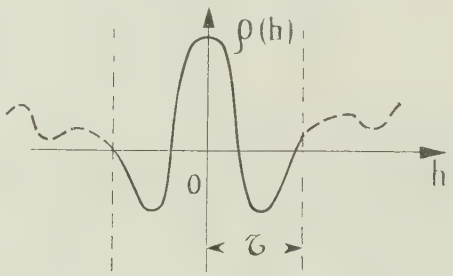


FIG. 33.

Un calculateur pourrait alors effectuer des mixages des amplitudes des $y_i\ (t)$ et des $\mathfrak{C}_i\ (t)$, arrivant en phase sur les traces, et conserver en mémoire les valeurs extrêmes résultantes qui seraient indiquées comme des réflexions possibles. Cette méthode a deux principaux défauts : une réflexion ne revenant pas en phase du fait du pendage de la couche réfléchissante serait supprimée, à moins qu'on ne procède à des mixages selon divers plans de pendage pour tester la présence de couches réfléchissantes à pendage. Le calculateur tiendrait très peu compte de la forme de la réflexion et il signalerait une foule de réflexions qu'un « pointeur » néglige du fait de son expérience sur les formations géologiques.

Manuscrit reçu en janvier 1961.

BIBLIOGRAPHIE

(1) AGARD et GRAU. — Les problèmes de la détection en prospection sismique. *Conférence au Centre Détection du Comité d'Action Scientifique de la Défense Nationale* (*C. A. S. D. N.*). 22 avr. 1959, *Inst. Franç. du Pétrole.*

(2) AGARD et GRAU. — Les problèmes de la détection en prospection sismique. *Conférence au Centre Détection du C. A. S. D. N.*, 31 mai 1960, *Inst. Franç. du Pétrole.*

(3) AGARD. — Thèse sur l'application de la loi log-

normale à la recherche minière. *Institut de Statistique de l'Université de Paris*, 1956.

(4) AGARD. — Problèmes d'attente dans une Compagnie aérienne. *Revue de la Société Française de Recherche Opérationnelle*, 1960, n° 17.

(5) BETHOUX. — Thèse de doctorat (*à paraître*).

(6) *Biometrika*. Tables for Statisticians, **I**.

(7) BLANC et LAPIERRE FORTET. — Théorie des fonctions aléatoires, *Masson*, Paris.

(8) BOIS, CHAUVEAU, GRAU et LAVERGNE. — Sismogrammes synthétiques. *Geophysical Prospecting*, 1960, **VIII**, n° 2.

(9) BURRAU. — The half invariants of two typical laws of errors with an application to the problem of dissecting a frequency curve into components. *Skand. Aktuarietidschrift*, 17-1934-16. *Voir aussi* (28).

(10) CRAWFORD, DOTY et LEE. — Continuous signal seismograph. *Geophysics*, **XXV**, 1, 95-105.

(11) DIX (C. H.). — Seismic prospecting for oil. *Harper*, New York.

(12) DYK (K.). — A comparison of additive and multiplicative compounding. *Geophysics*, 1956, **XXI**, n° 2.

(13) FORTET — Calcul des Probabilités. *Éd. du C. N. R. S.*

(14) FORTET. — Sur la détermination du spectre de l'inverse d'une fonction aléatoire et ses applications. *Publ. I. S. U. P.*, 1957, **VI**, fasc. 3.

(15) .FORTET. — Hypothesis testing and estimations for random Laplacian functions. *4e symposium de Berkeley*, 1960.

(16) A. HANEN. — Quelques problèmes de tests d'hypothèse et d'estimation en théorie des communications. *C. R. Ac. Sci.*, 1960, **250**, p. 3940-3942.

Étude géométrique du maximum de vraisemblance pour des processus stochastiques Laplaciens. *C. R. Ac. Sci.*, 1960, **250**, p. 4100-4101.

Propriétés du maximum de vraisemblance pour des processus Laplaciens. *C. R. Ac. Sci.*, 1960, **250**, p. 4268-4270.

(17) HORTON. — The structure of the noise back ground of a seismogram. *Geophysics*, Jul. 1955, **XX**, n° 3.

(18) LINDSEY J. P. — Elimination of seismic ghost reflections by means of a linear filter. *Geophysics*, 1960, **XXV**, n° 1.

(19) LOÈVE. — Processus stochastiques et mouvement brownien de Lévy. *Note*.

(19 *bis*). MANN H. B. — A theory of estimation for fundamental random process and Ornstein Uhlenbeck process. *Sankhya*, 1954, **13**, n° 4.

(20) MASSÉ L. — Probability distribution of reflections on seismic records. *Geophysics*, avr. 1955, **XX**, n° 2.

(21) NAVRATIL J. — Determination of the parameters of a compound distribution. *Pokroky Mat. Fys. Astr.*, **3**, 1958.

(22) PARZEN. — Regression analysis of continuous parameter time series. *Technical report* N° *33*, 1960, *Stanford University*.

(23) RICHARD-RIMBAUT. — Mesure du bruit et sismos multiples. *Seventh meeting of the European Association of Exploration Geophysicists*, The Hague, 8-10 déc. 1954.

(24) RICE. — Mathematical analysis of random noise. *Bell System Technical Journal*, 1944, 23.

(25) ROBINSON. — Predictive decomposition of time series with applications to seismic exploration, 1954. *Institute of Technology*, Cambridge, Massachussets.

(26) SEVASTYANOV. — Ergodic theorem for Markov process and its application to telephone systems with refusals. *Teoriya Veroyatnostei i ee Primeneniya*, 1957.

(27) STRIEBEL (C.). — Densities for Stochastic process. *Annals of Math Stat.*, 1959, **30**, p. 559-576.

(28) STROMGEN. — Tables and diagrams for dissecting a frequency curve into components. *Skand Aktuarietidschrift*, 1934, 17.7.54.

(29) ZAREMBA et LOMMICKI. — On estimating the spectral density function of a stochastic process. *J. Roy. Stat. Soc.*, 1957, **19**, n° 1.

DEUXIÈME THÈSE

PROPOSITIONS DONNÉES PAR LA FACULTÉ

PROBLÈMES D'ATTENTE AVEC DIVERSES RÈGLES DE PRIORITÉ

Vu et approuvé :

Paris, le 17 mars 1961

Le Doyen
de la Faculté des Sciences,

J. PÉRÈS

Vu et permis d'imprimer :

Le Recteur de l'Académie
de Paris,

J. SARRAILH

IMPRIMERIE PROTAT FRÈRES, MACON. Mars 1961 — DÉPOT LÉGAL 2e TRIMESTRE 1961
N° D'ORDRE CHEZ L'IMPRIMEUR : 7056 — N° D'ORDRE CHEZ L'ÉDITEUR : 244